



Europäisch s Patentamt
European Patent Office
Office européen des brevets



Publication number: **0 677 585 A1**

12

EUROPEAN PATENT APPLICATION

21 Application number: 95104393.4

22 Date of filing: 24.03.95

51 Int. Cl.⁶: **C12N 15/55, C12P 41/00, C12N 1/21, //(C12N1/21, C12R1:19),(C12P41/00, C12R1:19)**

A request for correction, here inclusion of page 2 of the description omitted upon filing, has been filed pursuant to Rule 88 EPC. A decision on the request will be taken during the proceedings before the Examining Division (Guidelines for Examination in the EPO, A-V, 3.).

30 Priority: 15.04.94 IT MI940726

43 Date of publication of application: 18.10.95 Bulletin 95/42

84 Designated Contracting States:
AT BE CH DE DK ES FR GB GR IE LI NL PT SE

71 Applicant: **ENIRICERCHE S.p.A.**
Via F. Maritano, 26
I-20097 S. Donato Milanese (Milano) (IT)

72 Inventor: **Grifantini, Renata**
Via Pietro Coletta 14
Milano (IT)
Inventor: **Frascotti, Gianni**
Via Gignous, 11
Milano (IT)
Inventor: **Galli, Giuliano**
Via Ferrandina, 14/A
San Donato Milanese (MI) (IT)
Inventor: **Grandi, Guido**
Nona Strada, 4
Segrate (San Felice) (MI) (IT)

74 Representative: **Gennari, Marco**
Eniricerche S.p.A.,
BREL,
Via F. Maritano, 26
I-20097 San Donato Milanese (MI) (IT)

54 **Process for the production of D-alpha-amino acids.**

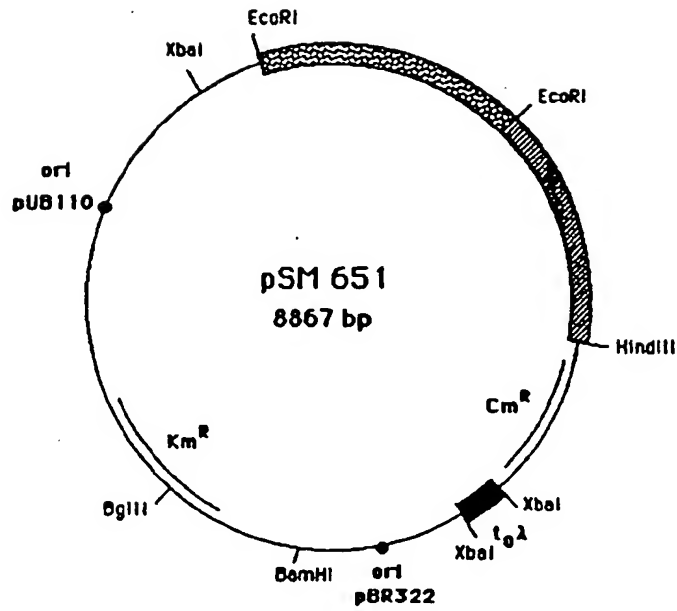
57 A process is described for the production of D- α -amino acids by the stereospecific conversion of racemic mixtures of 5-substituted hydantoins, where said conversion is carried out in the presence of a microorganism transformed with a plasmid capable of expressing at high levels and without inducers an enzymatic system capable of converting said hydantoins in the corresponding D- α -amino acids.

A plasmid is also described comprising the genes which encode said enzymatic system and a microorganism selected from *Escherichia coli* or *Bacillus subtilis* transformed with said plasmid.

D- α -amino acids are intermediates useful in the preparation of pharmacologically active substances, pesticides and sweeteners.

EP 0 677 585 A1

FIG. 3



The present invention relates to a process for the production of D- α -amino acids by the stereospecific conversion of racemic mixtures of 5-substituted hydantoin with a microorganism transformed with a plasmid capable of expressing in high yields and without inducers an enzymatic system capable of directly converting said hydantoin into the corresponding D- α -amino acids.

5 The term enzymatic system refers to a system consisting of D-hydantoinase and D-N-carbamylase enzymes.

D- α -amino acids are extremely valuable compounds useful for the preparation of pharmacologically active substances (for example, D-phenylglycine and D-parahydroxyphenylglycine are used in the synthesis of penicillins and cephalosporins), pesticides (D-valine for the synthesis of the insecticide fluvanilate) or
10 sweeteners (D-alanine).

The preparation of these enzymatic systems, however, requires the use of efficient inducers capable of stimulating the production of these enzymes on the part of the microorganisms. It is, in fact, known that the expression level of the enzymes D-hydantoinase and D-N-carbamylase is constitutively very low (Syldatk et al. (1990), "Advances in Biochem. Engineering/Biotechnology (Fiechter, A. Ed.), 41, pages 29-75, Springer-
15 Verlag, Berlin).

The inducers normally used are derivatives of hydantoin or nitrogenated cyclic compounds which are however easily metabolized by the microorganisms, or compounds such as uracil or thio-2-uracil or thymine which are not metabolized (Meyer et al., (1993), Fems Microbiol. Letters, 109: 67-74).

The use of inducers creates a series of drawbacks among which an increase in the production costs and a certain variability in the production yields of the enzymes. In addition, the expression level which can
20 be obtained in most of the microorganisms following induction is insufficient for economical use in industrial processes (Syldatk et al. (1987), Biotechnol. Lett., 9: 25-30; Yokozeki et al. (1987) Agric. Biol. Chem., 51, 715-722).

Recently the genes which encode the enzymes D-hydantoinase and D-N-carbamylase have been
25 individually sequenced and cloned (US 4.912.044 and EP-515-698).

More specifically, patent US 4.912.044 describes the preparation of D-hydantoinase by the fermentation of a microorganism transformed with a hybrid vector containing the hydantoinase gene whose expression is induced by temperature variation. The enzyme thus obtained is used for the production of D-N-carbamyl derivatives from 5-substituted hydantoin.

30 Patent application EP-515.698 describes, on the other hand, the preparation of D-N-carbamylase by the fermentation of a microorganism transformed with a plasmid comprising the carbamylase gene whose expression is chemically induced with IPTG. The enzyme thus obtained is used for the production of D- α -amino acids from N-carbamyl derivatives.

As industrial interest is directed towards the conversion of racemic hydantoin to D- α -amino acids, the
35 fact that the two enzymes are expressed in different strains involves the use of both and consequently the development of a process starting from two distinct fermentative processes.

This obviously increases the production costs and reduces the conversion kinetics. In fact, in order to complete the enzymatic reaction, the N-carbamyl derivative produced by the transformed microorganism containing the hydantoinase must pass through the bacterial membrane, spread into the reaction medium
40 and then proceed in the opposite direction to reach the second enzyme (carbamylase) present in the other strain. All this is particularly penalizing from the point of view of kinetics considering the reduced permeability of the bacterial membranes to the carbamyl derivatives (Olivieri et al. (1981), Biotechnol. Bioeng., 23, 2173-2183) and the inevitable dilution of the carbamyl itself in the reaction mixture.

Finally, the use of a double volume of biomass has a negative influence on the yields and degree of
45 purity of the final product.

In addition, the necessity of having to induce the expression of these enzymes creates a further problem thus making these processes of little interest for practical use.

The object of the present invention is to overcome the disadvantages of the known art described above.

In particular it has now been found, in accordance with the present invention, that the use of a particular
50 plasmid which contains the genes of D-hydantoinase and N-carbamylase put under the control of an appropriate synthetic promoter, enables the high expression of these enzymes to be obtained without inducers.

It is therefore possible to prepare a single microorganism transformed with said plasmid containing the two enzymatic activities inside. This solution solves not only the problems relating to kinetics due to the
55 limited permeability, as the two reactions occur inside the same cell where the concentration of the substrates is excellent, but also those relating to the requirement of inducers and treatment of the product and of the waste products.

In accordance with this, a first aspect of the present invention relates to a process for the production of D- α -amino acids by the stereospecific conversion of racemic mixtures of 5-substituted hydantoins characterized in that, the conversion reaction is carried out in the presence of a microorganism transformed with a plasmid capable of expressing at high levels and without inducers an enzymatic system capable of converting said hydantoins into the corresponding D- α -amino acids.

A further object of the present invention is the plasmid pSM651 comprising the genes which encode the enzymatic system.

Yet another object of the present invention is a microorganism transformed with the plasmid pSM651 capable of expressing with high efficiency and without inducers an enzymatic system capable of stereospecifically converting racemic mixtures of 5-substituted hydantoins into the corresponding D- α -amino acids.

A further object of the present invention relates to the use of said microorganisms or enzymatic system isolated from said microorganisms for the production of D- α -amino acids by the stereospecific conversion of racemic mixtures of 5-substituted hydantoins.

Further objects of the present invention will be evident from the description and examples below.

Brief description of the figures

Figure 1: Map of the plasmid pSM637 containing the carbamylase gene

Figure 2: Map of the plasmid pSM650 containing the hydantoinase gene

Figure 3: Map of the plasmid pSM651 containing the hydantoinase-carbamylase operon.

Figure 4 A-B: Nucleotide and amino acid sequence of carbamylase.

Figure 5 A-C: Nucleotide and amino acid sequence of hydantoinase.

Figure 6: SDS-PAGE (A) and Western-Blot (B) of the total proteins extracted from cultures of *E.coli* and *B.subtilis* transformed with the plasmid pSM651 wherein:

line 1: standard hydantoinase

line 2: standard carbamylase

line 3: *E.coli* (pSM671) control

line 4: *E.coli* SMC305

line 5: *B.subtilis* (pSM671) control

line 6: *B.subtilis* SMS373

The genes which encode the D-hydantoinase and D-N-carbamylase enzymes can be isolated from microorganisms such as *Pseudomonas*, *Hansenula*, *Agrobacterium*, *Aerobacter*, *Aeromonas*, *Bacillus*, *Moraxella*, *Brevibacterium*, *Flavobacterium*, *Serratia*, *Micrococcus*, *Arthrobacter* or *Paracoccus*. Specific examples of these microorganisms are *Bacillus macroides* ATCC 12905, *Aerobacter cloacae* IAM 1221, *Agrobacterium sp.* IP 1-671, *Agrobacterium radiobacter* NRRLB 11291, *Pseudomonas sp.* FERM BP 1900.

The isolation of the genes which encode the D-hydantoinase and D-N-carbamylase enzymes can be carried out by the construction of a gene library, representing the genome of the microorganism, identification of the clones containing the genes which encode said enzymes, analysis of the gene sequence, insertion of said genes into a vector and control of their expression.

The term gene library or genome bank means the combination of clones of a given host microorganism each of which carries a fragment of the chromosomal DNA deriving from the donor organism of which the bank is to be obtained. A bank is defined as being representative when the combination of the single fragments contained in each clone forms the majority of the chromosomal DNA of the donor organism.

According to a preferred embodiment of the process of the present invention, the strain *A.radiobacter* NRRL B-11291 is used as donor organism for the isolation of the genes which encode D-hydantoinase and D-N-carbamylase.

In practice, two genome banks of said microorganism are constructed in *E.coli* by digesting the chromosomal DNA separately with the restriction enzymes BamHI and SacI. Among the fragments obtained with the two digestions, those having dimensions normally of between 3,000 and 4,500 bp are purified. The selection is carried out by estimating the molecular weight of the D-hydantoinase and D-N-carbamylase enzymes of 50,000 and 34,000 Daltons respectively.

The two populations of BamHI and SacI fragments are then ligated to a vector of *E.coli* under such conditions as to facilitate the condensation of a single fragment to each molecule of the vector. The two ligase mixtures are used to transform cells of *E.coli* made competent as shown for example by Dagert, M. and Ehrlich (1979), (Gene, 6:23).

The two populations of colonies (genome banks) thus obtained, each of which carrying a hybrid plasmid i.e. consisting of the molecule of the vector and a chromosomal DNA fragment of *A.radiobacter*, are then

selected to identify those clones containing the hydantoinase and carbamylase genes.

The identification can be carried out by direct expression or using specific probes. The second method is preferably used. For the selection of the probes, in the case of carbamylase, reference was made to the knowledge of the amino-end sequence of carbamylase by *Comomonas* sp. 5222c (Ogawa et al. (1993), Eur.

5 J. Biochem., 212: 685-691).

On the basis of this sequence small oligonucleotides are synthesized which, once marked, are used for the screening of the genothecas by hybridization techniques (Maniatis et al., (1982), "Molecular Cloning: a laboratory manual", Cold Spring Harbor Laboratory).

This permitted the identification of a clone carrying a hybrid plasmid carrying a BamHI fragment
10 containing the nucleotidic sequence which encodes for the whole carbamylase. Analysis of said plasmid showed, in addition, the presence of a second incomplete ORF, placed on the other strand with respect to the carbamylase gene, which showed a homology with urease portions isolated from various microorgan-

isms. As ureases, like hydantoiasés, are enzymes belonging to the group of amido-hydrolases, it was
15 assumed that the incomplete ORF corresponded to that of hydantoise. This assumption was then confirmed by the enzymatic activity tests carried out on cellular extracts of cells carrying the identified gene.

In order to isolate the whole nucleotide sequence encoding hydantoinase, a screening of the gene library of the DNA of *A.radiobacter* digested with SacI was carried out by hybridization with an
oligonucleotide synthesized on the basis of the nucleotide sequence of the incomplete ORF.

20 The screening led to the isolation of a clone containing the whole hydantoinase gene. The genes thus isolated were sequenced using the sequenase version Kit 2.0 sold by United State Biochemical.

For the construction of a plasmid comprising both of the isolated genes vectors selected from plasmids, cosmids and bacteriophages known in the art, can be used.

The bifunctional plasmid of *E.coli* and *B.subtilis*, pSM671 CBS 205.94, is preferably used.

25 This plasmid, which comprises the genes which encode for resistance to kanamycin and chloramphenicol and has replication origins operable in *E.coli* and *B.subtilis*, is characterized in that it contains a synthetic promoter capable to direct with high efficiency and without inducers, the expression of the genes put under its control.

In practice, the DNA fragments containing the genes which encode the D-hydantoinase and D-N-
30 carbamylase enzymes are cloned into the plasmid pSM671 in the unique restriction sites EcoRI and HindIII obtaining the recombinant plasmid pSM651.

The construction can be carried out operating according to the general techniques known in the field of recombinant DNA. In order to verify whether these enzymes are expressed from *B.subtilis* and *E.coli*, cells transformed with said plasmid are cultured in a suitable culture medium. The total proteins, extracted from
35 the cellular lysate, analyzed on polyacrylamide gel showed the presence of two proteins having a molecular weight corresponding to that of the two enzymes; these proteins represent about 10% of the total proteins. These results confirm the capacity of *B.subtilis* and *E.coli* to express said enzymes with high efficiency and without inducers.

The enzymatic system of the present invention can be obtained by culturing the strains *E.coli* or
40 *B.subtilis* transformed with the plasmid pSM651, under aerobic conditions, in an aqueous medium containing assimilable sources of carbon and nitrogen as well as various cations, anions and, possibly, traces of vitamins, such as biotin, thiamine, or amino acids.

Assimilable carbon sources comprise carbohydrates such as glucose, hydrolized amides, molasses, sucrose or other conventional carbon sources.

45 Examples of nitrogen sources can be selected, for example, from mineral ammonium salts, such as ammonium nitrate, ammonium sulphate, ammonium chloride or ammonium carbonate and urea or materials containing organic or inorganic nitrogen such as peptone, yeast extract or meat extract.

The following cations and anions are equally suitable for the object of the present invention: potassium, sodium, magnesium, iron, calcium, acid phosphates, sulphates, chlorides, manganese, and nitrates.

50 The fermentation is carried out, under stirring, at a temperature of between 25° and 40° C, preferably between 30° and 37° C and at a pH of between 6 and 7.5, preferably between 6.5 and 7.0.

The cells (biomass) recovered from the culture medium by means of the conventional techniques such as centrifugation or filtration are used in the conversion phase of the racemic mixtures of 5-substituted
hydantoin.

55 Alternatively, the conversion reaction can be carried out using either the cellular xtract obtained from the disintegration of the cells by sonication or French-Press, or enzymes purified or partially purified with the conventional methods, or enzymes immobilized on insoluble supports.

Numerous hydantoin substituted in position 5 can be used in the process of the present invention. Possible substituents in position 5 are selected from a linear or branched alkyl group with a number of carbon atoms of between 1 and 6, which can be mono or polysubstituted with hydroxy, carboxy, hydrosulphide or amino groups or a phenyl or benzyl group which, in turn, can contain one or more substituents in ortho, meta and para position. Examples of 5-substituted hydantoins are: D,L-5-phenylhydantoin, D,L-5-para-hydroxyphenylhydantoin, D,L-5-methylhydantoin, D,L-5-isopropylhydantoin, D,L-5-thienylhydantoin, D,L-5-para-methoxyphenylhydantoin, D,L-5-para-chloro phenylhydantoin, D,L-5-benzylhydantoin.

The conversion of the hydantoins into the corresponding D- α -amino acids is carried out in a nitrogen atmosphere in a hermetically closed apparatus, at a temperature of between 20 and 60 °C, preferably between 30 and 45 °C.

The pH of the reaction medium is maintained within values of between 6 and 10 and preferably between 7 and 8.5. This regulation of the pH can be carried out, for example, by adding a base aqueous solution such as an aqueous solution of ammonia, potassium hydroxide, sodium hydroxide, sodium or potassium carbonate.

The initial concentration of the hydantoins is generally between 2% and 30% by weight. As a result of the stereospecificity of the enzymes produced from the strains of the present invention, only the D-enantiomorphs of the hydantoins are hydrolyzed. As hydantoins however, spontaneously racemize more or less rapidly under the operating conditions, the L-enantiomorphs are completely converted into the corresponding D- α -amino acids.

The quantity of biomass which is added to the reaction mixture depends on the particular affinity of the substrate towards the enzymes. Generally a ratio by weight biomass/hydantoins of between 1/1 and 1/50 can be used.

When the conversion reaction is carried out under optimum conditions a yield of 95-98% is obtained.

The D- α -amino acids prepared with the process of the present invention can be recovered from the reaction medium with the conventional methods such as ion-exchange chromatography or precipitation of the amino acid at its isoelectric point.

The plasmid pSM651 was deposited at the Bureau Voor Schimmelcultures, SK Baarn (Holland) as *E.coli* SMC305 where it received the deposit number CBS 203.94.

The following experimental examples provide a better illustration of the present invention but do not limit it in any way.

Example 1

Extraction of the chromosomal DNA from A.radiobacter

100 ml of fermentation medium having the following composition:

1% glucose, 0.3% yeast extract, 1.36% KH_2PO_4 , 0.02% $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ (pH 7.0) were inoculated with the strain *A.radiobacter* (NRRLB 11291) and maintained under stirring (220 rpm) at 30 °C for 24 hours.

The cells were then recovered by centrifugation of the culture broth in an SS34 rotor model Sorvall RC-5B (at 4 °C and 5000 rpm for 10 minutes) and then washed (2x120 ml) with a solution (TE) containing 1 mM EDTA, 10 mM Tris-HCl, pH 7.4. The resulting suspension was centrifuged again as above and the cells were recovered and resuspended in 9.5 ml of TE solution. After adding 0.5 ml of 10% SDS (sodium dodecylsulphate) and 50 μl of a solution of Proteinase K (20 mg/ml), the suspension was incubated at 37 °C for 1 hour.

1.8 ml of NaCl 5 M and 1.5 ml of a solution consisting of 10% hexadecyltrimethyl ammonium bromide (CTAB) in 0.7 M NaCl were subsequently added and the resulting solution was incubated at 65 °C for 20 minutes. The solution was then deproteinized with an equal volume of chloroform/isoamyl alcohol (24/1, v/v) and the DNA was precipitated with 0.6 volumes of isopropanol. The DNA was washed with 1 ml of ethanol (70%) and recovered with a glass rod. The recovered DNA was finally dissolved in 4 ml of TE and its concentration was determined by spectrophotometry at 260 nm.

The chromosomal DNA was purified again by centrifugation on a gradient of CsCl (1%) containing 0.1 mg/ml of ethidium bromide (55,000 rpm for 16 hours in a Beckman rotor V65Ti).

The DNA band was visualized under a UV light and the ethidium bromide was removed by extraction with butanol saturated in H_2O . After dialysis against a TE buffer, the DNA was precipitated with ethanol and resuspended in the desired volume.

Example 2

Construction of a genomic bank of A. radiobacter

5 Aliquots (10 µg) of the DNA thus obtained were digested, separately, with 25 units of each of the restriction enzymes EcoRI, PstI, BamHI, SacI, and SphI (Boehringer) operating according to the instructions of the producer.

After blocking the enzymatic reactions at 65 °C for 10 minutes, the reaction mixtures were charged onto agar gel at 0.8% and run at 100 volts for 2 hours. The DNA bands, visualized by coloring with EtBr (0.5 gamma/ml), were then transferred onto a nylon filter (Boehringer) and after lysis with NaOH, the DNA was
10 immobilized according to the Southern blot technique (Maniatis et al., "Molecular Cloning: a practical laboratory manual", Cold Spring Harbor, New York, 1982).

The filter was hybridized at 45 °C with each of the degenerated oligonucleotides, conceived on the basis of the amino-end of the carbamylase of Comamonas sp. E222c (Ogawa et al., (1993), Eur. J. Biochem., 212: 685-691), having the sequence:
15

1) 5'CGA ATT GTA AAT TAT GCA GCA GC 3'

20 A G C G C C G G
 C A C C C
 T T T T
25

2) 5'GGA CCA ATT CAA CGA GC 3'

 G G C G G
30 C C A C
 T T T
35

3) 5'CGA GCA GAT GTA ATG GA 3'

 A G G C G
40 C C C
 T T T

45 These oligonucleotides were synthesized using the automatic System OLIGO 1000 system of Beckmann and then marked at the 5'end using the kit DIG SYSTEM (Boehringer). The hybridization reaction with probe 2 gave positive signals. In particular, the DNA digested with BamHI generated a fragment of about 4000 bp capable of hybridizing the probes.

To isolate the BamHI fragment thus identified, 10 µg of chromosomal DNA were suspended in 50 µl of
50 buffer 10 mM Tris-HCl pH 8, 100 mM NaCl, 5 mM MgCl₂, 1 mM 2-mercaptoethanol and incubated at 37 °C for 4-5 hours in the presence of 25 U of the enzyme BamHI.

The digestion mixture was then subjected to electrophoresis on agar gel at 0.8% and, after colouring with EtBr, DNA fragments of 3,500 - 4,500 bp were electroeluted in the electrophoresis buffer (Maniatis et al. "Molecular Cloning: a practical laboratory manual", Cold Spring Harbor, New York 1982).

55 The chromosomal DNA fragments in the plasmid pUC18 (BRL) were then cloned. In practice, 20 ng of this plasmid, previously linearized with the restriction enzyme BamHI, were ligated with 100 ng of the chromosomal DNA fragments in 20 µl of mixture containing 66 mM Tris-HCl pH 7.6, 1 mM ATP, 10 mM MgCl₂, 10 mM Dithiothreitol (DTT), in the presence of 1 U of T4 DNA ligase, at 16 °C for a night.

The ligase mixture was used for transforming cells of *E.coli* JM101 (BRL) made competent with 50 mM CaCl₂ (Dagert, M. and Ehrlich (1979), Gene, 6:23).

The transformants were subsequently selected on plates of LB medium (8 g/l Bactotryptone (DIFCO), 5 g/l NaCl, 15 g/l Agar (DIFCO), 0.5 g/l yeast extract) to which 40 µg/ml of X-Gal (5-bromo-4-chloro-3-indolyl-D-thio-galactopyranoside) and 100 µg/ml of ampicillin had been added.

Operating as described above numerous positive recombinant colonies (white) were obtained which were easily distinguishable from those not recombinant (blue).

The positive clones were transferred onto nylon filters (Boehringer) and the DNA extracted from these clones was hybridized under the same conditions using probe 2 which had responded positively to hybridization with the chromosomal DNA.

The plasmids extracted from the clones which gave a positive signal were sequenced using the Sequenase version 2.0 Kit (United States Biochemical). One of these plasmids, containing the complete carbamylase gene (915 bp) was called pSM652.

Figure 4 shows the nucleotidic and amino acidic sequence of carbamylase.

Example 3

Isolation of the hydantoinase gene of *A.radiobacter*

Analysis of the plasmid pSM652 showed the presence of a second incomplete ORF, situated on the other strand with respect to the carbamylase gene, which showed a homology with urease portions isolated from various microorganisms.

As ureases, like hydantoinases, are enzymes belonging to the group of amidohydrolases, it was assumed that the incomplete ORF corresponded to that of hydantoinase. The assumption was then confirmed by enzymatic activity tests carried out on cellular extracts of cells carrying the identified gene.

In order to isolate the whole nucleotidic sequence encoding the hydantoinase, the same Southern Blot used for isolating the carbamylase was hybridized using as probe the oligonucleotide having the sequence: 5' ATC GTA ACC GCG GAG GGG ATT TCT CCC 3'

This oligonucleotide, homologous to the 5'end region of the nucleotidic sequence of identified partial ORF, was synthesized and marked as shown in example 2. Among the positive bands for this probe a band of about 3500 bp obtained from the digestion of the DNA with the enzyme SacI, was identified.

Operating as shown in example 2 a genomic bank of chromosomal DNA of *A.radiobacter* digested with SacI was then constructed. Screening of this bank led to the isolation of the plasmid pSM653 containing the whole gene for hydantoinase whose nucleotide and amino acid sequence is shown in figure 5.

Example 4

Cloning of the carbamylase gene

1) Amplification of the carbamylase gene

The plasmid pSM652 was amplified by the Polymerase Chain Reaction (PCR) technique, according to the indications specified by Leung et al. (Leung D.W., Chen E., Goeddel D.V., Technique - a journal of methods in cell and molecular biology, 1, No. 1 (1989): pages 11-15), using the pair of oligonucleotides:

(1) 5' GGG AAT TCT TAT GAC ACG TCA G 3' (FORWARD)

EcoRI

(2) 5' CCC AAG CTT CAA AAT TCC GCG AT 3' (REVERSE)

HindIII

The oligonucleotide (2) also allowed the restriction site EcoRI present inside the carbamylase gene near 3'end, to be eliminated.

The amplification was carried out in a DNA Thermal Cycler 480 apparatus (Perkin - Elmer Cetus) using a reaction mixture (100 µl) containing 10 mM Tris HCl pH 8.3, 1.5 mM MgCl₂, 50 mM KCl, 0.01% (weight/volume) of gelatine, 1 ng of pSM652, 1 µM of the two primers, 200 µM of dNTP, 0.5 Units of Taq polymerase (Perkin Elmer).

- 5 A drop of mineral oil is added and the mixture is denatured for 4 minutes at 94 °C and the cyclic program is started, which comprises:
1 minute at 94 °C (denaturation)
1 minute at 55 °C (annealing)
2 minutes at 72 °C (elongation)
- 10 for a total of 30 cycles, followed by 8 minutes at 72 °C (final extension).

- The amplification product thus obtained was treated with phenol-chloroform (1:1), precipitated with NaCl (1/10 vol/vol) and EtOH (2 vol) and resuspended in 20 µl of H₂O. After cutting with the restriction enzymes EcoRI and HindIII (5 U) suitable for cloning into the plasmid pSM671 (CBS 205.94) the DNA fragments were purified on low-melting gel (SeaPlaque, FMC BioProducts) at 1.0% and the bands eluted by the gel were
- 15 treated with GEIase (Epicentre Technologies) (1 U every 300 µg of gel weighed) for 1.5 hours at 45 °C.

At the same time, 50 ng of the plasmid pSM671 were cut with the same restriction enzymes.

- The plasmid and fragments were ligated in 10 µl of reaction mixture (DNA 20 ng/ml) and 2 µl of this mixture were used for transforming cells of *E.coli* 71/18 made competent with CaCl₂ (Dagert and Ehrlich, Gene, 6: 23, 1979). The transformants were selected on plates of LB medium containing 20 µg/ml of chloramphenicol.
- 20

The plasmid DNA extracted from the positive clones was analyzed to verify exact insertion into the carbamylase gene and the absence of possible errors caused by the amplification.

One of these plasmids was called pSM637.

The strain of *E.coli* containing the plasmid pSM637 was called SMC307.

- 25 Cells of *B.subtilis* SMS108 NRRLB-15.898 made competent as described in "Molecular Biology Methods for Bacillus". (1990) (Harwood and Cutting (eds) Wiley and Sons) were transformed with 100 ng of the plasmid pSM637 operating according to the known techniques, and the transformed strain was called SMS374.

30 Example 5

Expression of the carbamylase gene in *E.coli* and *B.subtilis*

- The object of the experiment was to verify the ability of the transformed strains (*E.coli* SMC307 and
- 35 *B.subtilis* SMS 374) to express the carbamylase gene without inducers.

A preculture on slant of the strain *E.coli* SMS307 and *B.subtilis* SMS 374 was inoculated into two 100 ml flasks containing, respectively, 10 ml of LB medium to which 20 µg/ml of chloramphenicol had been added and 10 ml of VY medium to which 5 µg/ml of chloramphenicol had been added. The flasks were incubated, under stirring, (220 rpm), at 37 °C for 16 hours.

- 40 The cells were recovered by centrifugation (12,000 rpm, 4 °C, for 1 minute) of the two culture broths, resuspended in 300 µl of buffer 20 mM Tris-HCl pH 7.5, 20 mM BMeOH, 20% glycerol and lysed by sonication (Soniprep150, MSE 1 minute impulses, at average voltage). Aliquots (15 µl) of the two lysates were charged onto polyacrylamide gel at 10% and run at 20 mA for three hours. The proteic bands were visualized by colouring with Coomassie R-250 (Laemmli, Nature: 227, 680, 1970). After colouring with
- 45 Coomassie a proteic band was revealed with a molecular weight of 34,000 D absent in the extracts of untransformed strains *B.subtilis* SMS108 and *E.coli* 71/18. In addition, densitometric analysis carried out on the same gel coloured with Coomassie showed that this protein was expressed in both of the transformed strains as one of the prevalent proteins (10% with respect to the total proteins).

50 Example 6

Cloning of the hydantoinase gene

- The plasmid pSM653 (1 µg) was digested with the restriction enzymes EcoRV and Sall (4 U)
- 55 (Boehringer) at 37 °C for 1 hour.

The digestion mixture was then subjected to electrophoresis on agar gel at 0.8% (low melting) and, after colouring with EtBr, the DNA band corresponding to an EcoRV-Sall fragment of 1300 bp was recut and the DNA extracted with the Gelase TM method (EPICENTRE Technologies). As this fragment has a small

region missing at the 5' end and a portion of 70 bp at 3' end, the whole hydantoinase gene was reconstructed using two linkers having the sequence:

LINKER 5'

5' AATTCTTATG GAT 3'

EcoRI

LINKER 3'

5' TCGACGAGGG AACCTACGTG GGGGCGCCGA CGGATGGCCA

SalI

GTTCCGGAAG CGCCGCAAAT ACAAGCAATA AGGAGG 3'

EcoRI

40 ng of the 1300 bp fragment, 40 ng of the linker 3', 10 ng of the linker 5' and 50 ng of the plasmid pSM671 CBS 205.94, previously linearized with EcoRI, were then ligated in a ligase mixture containing 1 U of T4 DNA ligase, incubating at 12°C for 16 hours. The ligase mixture was subsequently used to transform competent cells of *E.coli* 71/18 and the transformants were selected on plates of LB medium to which 20 µg/ml of chloramphenicol had been added.

The plasmid DNAs isolated from some of the positive clones were analyzed to identify the clones containing the complete sequence of the hydantoinase gene.

One of these plasmids was called pSM650 and the strain of *E.coli* containing said plasmid was marked with the abbreviation SMC308.

100 ng of the plasmid pSM650 were used to transform competent cells of *B.subtilis* SMS108. The resulting strain was called SMS375.

Example 7

Expression of the hydantoinase gene in *E.coli* and *B.subtilis*

The object of the experiment was to verify the capacity of the transformed strains (*E.coli* SMS308 and *B.subtilis* SMS375) to express the hydantoinase gene without inducers.

A preculture on slant of the strain *E.coli* SMS308 and *B.subtilis* SMS375 was inoculated into two 50 ml flasks containing, respectively, 10 ml of LB medium to which 5 µg/ml of chloramphenicol had been added and 10 ml of VY medium to which 20 µg/ml of chloramphenicol had been added. The flasks were incubated, under gentle stirring, (220 rpm), at 37°C for 16 hours.

The cells were recovered by centrifugation (12,000 rpm, 4°C, for 1 minute) of the two culture broths, resuspended in 300 µl of buffer 20 mM Tris-HCl pH 7.5, 20 mM BMeOH, 20% glycerol and lysed by sonication (1 minute impulses, at average voltage). Aliquots (15 µl) of the two lysates were charged onto polyacrylamide gel at 10% and run at 20 mA for three hours. The proteic bands were visualized by colouring with Coomassie R-250 (Laemmli, Nature: 227, 680, 1970). After colouring with Coomassie a proteic band was revealed with a molecular weight of 50,000 Daltons absent in the extracts of untransformed strains *B.subtilis* SMS108 and *E.coli* 71/18. In addition, densitometric analysis carried out on the same gel coloured with Coomassie showed that this protein was expressed in the two transformed strains as one of the prevalent proteins (10% with respect to the total proteins).

Example 8

Cloning of the hydantoinase-carbamylase operon

5 The plasmid pSM650 (1 μ g) was digested with the enzyme EcoRI (5 U) at 37 °C for 1 hour. The EcoRI-EcoRI fragment of about 1380 bp containing the hydantoinase gene was purified by agar gel at 0.8% with the Gelase TM method. 20 ng of this fragment were ligated with 50 ng of the plasmid pSM637 linearized with EcoRI. The reaction was carried out in a ligase buffer containing 1 U of T4 DNA ligase, at 16 °C for 16 hours.

10 The ligase mixture was used to transform competent cells of E.coli 71/18.

The transformants were subsequently selected on plates of LB medium (8 g/l Bactotryptone (DIFCO), 5 g/l NaCl, 15 g/l Agar (DIFCO), 0.5 g/l yeast extract) to which 20 μ g/ml of Chloramphenicol had been added.

15 The positive clones were analyzed by restriction analysis to verify the correct insertion into the two genes. The plasmid containing the hydantoinase-carbamylase operon was called pSM651 and the strain of E.coli containing said plasmid was marked with the abbreviation SMC305.

Competent cells of B.subtilis SMS108 were transformed with 100 ng of this plasmid. One of the positive clones was called SMS373.

20 Example 9

Expression of the hydantoinase-carbamylase operon

25 E.coli SMS305 and B.subtilis SMS373 were cultured, respectively, in 100 ml of LB medium to which 20 μ g of chloramphenicol had been added and in 100 ml of VY medium to which 5 μ g of chloramphenicol had been added, at 37 °C for 16 hours, under stirring (200 rpm). The proteic extracts obtained from the cellular lysates were analyzed as described in example 7. The results showed the presence of two proteins corresponding to hydantoinase and carbamylase (figure 6). To evaluate the activity of these enzymes, a reaction kinetics was carried out using 20 mM (D,L) parahydroxyphenyl-hydantoin as substrate or alternatively 5-phenyl-hydantoin (in 200 mM of phosphate buffer pH 8) and following the conversion into the corresponding D- α -amino acid with the evolution of ammonia. The process adopted is described by Weatherburn, M.W., (1967), (Anal. Chem., 39:971).

Example 10

35

Conversion of D,L-5-phenylhydantoin to D-phenylglycine

40 A suspension of 2 g of D,L-5-phenyl-hydantoin in 100 ml of Na-phosphate 0.2 M buffer pH 8.0 was charged into an apparatus equipped with a stirrer and thermostat-regulated at 40 °C. After degassing with nitrogen at 40 °C for 5 minutes, 5 g (humid weight) of biomass was introduced, coming from a culture of E.coli SMS305, carried out as described in example 9.

45 After the apparatus had been hermetically closed, the reaction mixture was maintained under a nitrogen atmosphere, at 40 °C for 24 hours. Polarimetric and thin layer chromatographic analysis (J. of Chromatography, 80: 199-204), 1973) of an aliquot of the reaction mixture showed the complete hydrolysis of the starting substrate to D-phenylglycine.

50 After separation of the biomass by centrifugation of the reaction mixture at 6000 rpm for 10 minutes, the supernatant was acidified to pH 1.0 with HCl 6 M and charged onto a column (2.6 x 20 cm) of Amberlite IR 120 (activated with HCl). The column was then washed with water and eluted with an ammonia solution at 5% in water. The eluate was decoloured with decolouring carbon (C.Erba), and the decoloured solution was concentrated under vacuum and brought to pH 5.8. The crystals thus obtained were recovered by filtration and recrystallized from water. The white powder obtained (1.63 g) showed a specific rotation $[\alpha]_D^{20} = -156^\circ$ (c = 1, 1 N HCl). The IR spectrum was in agreement with that of the standard D-phenylglycine.

55

Example 11

Conversion of D,L-5-phenylhydantoin to D-phenylglycine.

5 The same procedure was carried out as in example 10, using 5 g (humid weight) of biomass coming from the culture of *E.coli* SMS305 and 10 g of D,L-5-phenylhydantoin in 100 ml of Na-phosphate 0.2 M buffer pH 8.0. The reaction was carried out under a nitrogen atmosphere, at 40 °C for 90 hours. The white powder obtained (8.1 g) showed a specific rotation $[\alpha]_D^{20} = -156.5^\circ$ ($c = 1, 1 \text{ N HCl}$). The IR spectrum agreed with that of the standard D-phenylglycine.

Example 12

Conversion of D,L-5-para-hydroxy-phenylhydantoin to D-para-hydroxy-phenylglycine

15 The same procedure was carried out as in example 10, using 2.5 g (humid weight) of biomass and 1 g of D,L-5-para-hydroxy-phenylhydantoin. The D-para-hydroxy-phenylglycine obtained as a white powder (0.82 g) showed a specific rotation $[\alpha]_D^{20} = -158^\circ$ ($c = 1, 1 \text{ N HCl}$). The IR spectrum was in agreement with that of the standard D-phenylglycine.

Example 13

Conversion of D,L-5-para-hydroxy-phenylhydantoin to D-para-hydroxy-phenylglycine

25 The same procedure was carried out as in example 10, using 2.5 g of biomass (humid weight) obtained from the culture of *E.coli* SMS305 and 8 g of D,L-5-para-hydroxy-phenylhydantoin.

The reaction was carried out under a nitrogen atmosphere, at 40 °C for 170 hours. The D-para-hydroxy-phenylglycine obtained as a white powder (6.6 g) showed a specific rotation $[\alpha]_D^{20} = -157.8^\circ$ ($c = 1, 1 \text{ N HCl}$). The IR spectrum was in agreement with that of the standard D-phenylglycine.

Example 14

Conversion of D,L-5-isopropylhydantoin to D-valine

35 The same procedure was carried out as in example 10, using 5.0 g of biomass (humid weight) obtained from the culture of *E.coli* SMS305 and 2 g of D,L-5-isopropylhydantoin.

The reaction was carried out under a nitrogen atmosphere, at 40 °C for 240 hours. The D-valine obtained as a white powder (0.8 g) showed a specific rotation $[\alpha]_D^{20} = -27.5^\circ$ ($c = 5, 6 \text{ N HCl}$). The IR spectrum agreed with that of the standard D-valine.

SEQUENCE LISTING

NUMBER OF SEQUENCES: 19

5 (1) INFORMATION FOR SEQ ID NO:1:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGHT: 23 base pairs
 10 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear
 15 (ii) MOLECULE TYPE: DNA (Genomic)
 (xi) SEQUENCE DESCRIPTION:
 CGAATTGTAA ATTATGCAGC AGC 23
 20 (1) INFORMATION FOR SEQ ID NO:2:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGHT: 23 base pairs
 25 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear
 30 (ii) MOLECULE TYPE: DNA(Genomic)
 (xi) SEQUENCE DESCRIPTION:
 AGGATCGTGA ACTACGCGGC GGC 23
 35 (1) INFORMATION FOR SEQ ID NO:3:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGHT: 23 base pairs
 40 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear
 45 (ii) MOLECULE TYPE: DNA(Genomic)
 (xi) SEQUENCE DESCRIPTION:
 CGCATAGTCA ATTATGCCGC CGC 23
 50 (1) INFORMATION FOR SEQ ID NO:4:

55

(i) SEQUENCE CHARACTERISTICS:

(A) LENGHT: 23 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA(Genomic)

(xi) SEQUENCE DESCRIPTION:

CGTATTGTTA ATTATGCTGC TGC

23

(1) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGHT: 17 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA(Genomic)

(xi) SEQUENCE DESCRIPTION:

GGACCAATTC AACGAGC

17

(1) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGHT: 17 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA(Genomic)

(xi) SEQUENCE DESCRIPTION:

GGGCCGATCC AGCGGGC

17

(1) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGHT: 17 base pairs

(B) TYPE: Nucleic acid

EP 0 677 585 A1

(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear
5 (ii) MOLECULE TYPE: DNA(Genomic)
(xi) SEQUENCE DESCRIPTION:
GGCCCCATAC AACGCGC 17
10 (1) INFORMATION FOR SEQ ID NO:8:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGHT: 17 base pairs
15 (B) TYPE: Nucleic acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear
20 (ii) MOLECULE TYPE: DNA(Genomic)
(xi) SEQUENCE DESCRIPTION:
GGTCCTATTC AACGTGC 17
25 (1) INFORMATION FOR SEQ ID NO:9:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGHT: 17 base pairs
30 (B) TYPE: Nucleic acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear
35 (ii) MOLECULE TYPE: DNA(Genomic)
(xi) SEQUENCE DESCRIPTION:
CGAGCAGATG TAATGGA 17
40 (1) INFORMATION FOR SEQ ID NO:10:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGHT: 17 base pairs
45 (B) TYPE: Nucleic acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear
50 (ii) MOLECULE TYPE: DNA(Genomic)

55

(x1) SEQUENCE DESCRIPTION:

AGGGCGGACG TGATGGA

17

(1) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGHT: 17 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA(Genomic)

(x1) SEQUENCE DESCRIPTION:

CGCGCCGATG TCATGGA

17

(1) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGHT: 17 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA(Genomic)

(x1) SEQUENCE DESCRIPTION:

CGTGCTGATG TTATGGA

17

(1) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGHT: 27 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA(Genomic)

(x1) SEQUENCE DESCRIPTION:

ATCGTAACCG CGGACGGGAT TTCTCCC

27

(1) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGHT: 22 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA (Genomic)

(ix) FEATURE:

(A) NAME: Primer

(xi) SEQUENCE DESCRIPTION:

GGGAATTCTT ATGACACGTC AG 22

(1) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGHT: 23 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA(Genomic)

(ix) FEATURE:

(A) NAME: Primer

(xi) SEQUENCE DESCRIPTION:

CCCAAGCTTC AAAATTCCGC GAT 23

(1) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGHT: 13 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA(Genomic)

(ix) FEATURE:

(A) NAME: Linker

(x1) SEQUENCE DESCRIPTION:

AATTCTTATG GAT 13

(1) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGHT: 76 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA(Genomic)

(ix) FEATURE:

(A) NAME: Linker

(x1) SEQUENCE DESCRIPTION:

TCGACGAGGG AACCTACGTG GGGGCGCCGA CGGATGGCCA 40

GTTCCGGAAG CGCCGCAAAT ACAAGCAATA AGGAGG 76

(1) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGHT: 915 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA (Genomic)

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Agrobacterium radiobacter

(x1) SEQUENCE DESCRIPTION:

ATG ACA CGT CAG ATG ATA CTT GCT GTC GGA CAG CAA GGC CCC ATC 45

Met Thr Arg Gln Met Ile Leu Ala Val Gly Gln Gln Gly Pro Ile

5

10

15

GCG CGA GCG GAG ACA CGC GAA CAG GTG GTT GGC CGC CTC CTC GAC 90

Ala Arg Ala Glu Thr Arg Glu Gln Val Val Gly Arg Leu Leu Asp

20

25

30

EP 0 677 585 A1

| | | |
|----|---|-----|
| | ATG TTG ACG AAC GCA GCC AGC CGG GGC GTG AAC TTC ATC GTC TTT | 135 |
| | Met Leu Thr Asn Ala Ala Ser Arg Gly Val Asn Phe Ile Val Phe | |
| 5 | 35 40 45 | |
| | CCC GAG CTT GCG CTC ACG ACC TTC TTC CCG CGC TGG CAT TTC ACC | 180 |
| | Pro Glu Leu Ala Leu Thr Thr Phe Phe Pro Arg Trp His Phe Thr | |
| 10 | 50 55 60 | |
| | GAC GAG GCC GAG CTC GAT AGC TTC TAT GAG ACC GAA ATG CCC GGC | 225 |
| | Asp Glu Ala Glu Leu Asp Ser Phe Tyr Glu Thr Glu Met Pro Gly | |
| 15 | 65 70 75 | |
| | CCG GTG GTC CGT CCA CTC TTT GAG ACG GCC GCC GAA CTC GGG ATC | 270 |
| | Pro Val Val Arg Pro Leu Phe Glu Thr Ala Ala Glu Leu Gly Ile | |
| 20 | 80 85 90 | |
| | GGC TTC AAT CTG GGC TAC GCC GAA CTC GTC GTC GAA GGC GGC GTC | 315 |
| | Gly Phe Asn Leu Gly Tyr Ala Glu Leu Val Val Glu Gly Gly Val | |
| 25 | 95 100 105 | |
| | AAG CGT CGC TTC AAC ACG TCC ATT CTG GTG GAT AAG TCA GGC AAG | 360 |
| | Lys Arg Arg Phe Asn Thr Ser Ile Leu Val Asp Lys Ser Gly Lys | |
| 30 | 110 115 120 | |
| | ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG GGT CAC AAG GAG | 415 |
| | Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro Gly Hys Lys Glu | |
| 35 | 125 130 135 | |
| | TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG CGT TAT TTC | 450 |
| | Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys Arg Tyr Phe | |
| 40 | 140 145 150 | |
| | GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC GCC GCG | 495 |
| | Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp Ala Ala | |
| 45 | 155 160 165 | |
| | AAA ATG GGG ATG TTC ATC TGC AAC GAT CGC CGC TGG CCT GAA ACG | 540 |
| | Lys Met Gly Met Phe Ile Cys Asn Asp Arg Arg Trp Pro Glu Thr | |
| 50 | 170 175 180 | |
| 55 | | |

EP 0 677 585 A1

TGG CGG GTG ATG GGA CTT AAG GGC GCC GAG ATC ATC TGC GGC GGC 585
 Trp Arg Val Met Gly Leu Lys Gly Ala Glu Ile Ile Cys Gly Gly
 5 185 190 195
 TAC AAC ACG CCG ACC CAC AAT CCC CCC GTT CCC CAG CAC GAC CAT 630
 Tyr Asn Thr Pro Thr His Asn Pro Pro Val Pro Gln His Asp His
 10 200 205 210
 CTG ACG TCC TTC CAC CAC CTT CTG TCG ATG CAG GCC GGG TCG TAC 675
 Leu Thr Ser Phe His His Leu Leu Ser Met Gln Ala Gly Ser Tyr
 15 215 220 225
 CAA AAC GGC GCC TGG TCC GCG GCG GCC GGC AAG GTC GGC ATG GAG 720
 Gln Asn Gly Ala Trp Ser Ala Ala Ala Gly Lys Val Gly Met Glu
 20 230 235 240
 GAG GGG TGC ATG CTG CTC GGC CAT TCG TGC ATC GTG GCG CCG ACC 765
 Glu Gly Cys Met Leu Leu Gly His Ser Cys Ile Val Ala Pro Thr
 25 245 250 255
 GGC GAA ATC GTT GCC CTG ACC ACG ACG TTG GAA GAC GAG GTG ATC 810
 Gly Glu Ile Val Ala Leu Thr Thr Thr Leu Glu Asp Glu Val Ile
 30 260 265 270
 ACC GCC GCC GTC GAT CTC GAC CGC TGC CGG GAA CTG CGC GAA CAC 855
 Thr Ala Ala Val Asp Leu Asp Arg Cys Arg Glu Leu Arg Glu His
 35 275 280 285
 ATC TTC AAT TTC AAA GCC CAT CGT CAG CCA CAG CAC TAC GGT CTG 900
 Ile Phe Asn Phe Lys Ala His Arg Gln Pro Gln His Tyr Gly Leu
 40 290 295 300
 ATC GCG GAA TTT TGA 915
 Ile Ala Glu Phe STOP
 45 (1) INFORMATION FOR SEQ ID NO:19:
 (1) SEQUENCE CHARACTERISTICS:
 (A) LENGHT: 1373 base pairs
 50 (B) TYPE: Nucleic acid
 55

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA (Genomic)

(vi) ORIGINAL SOURCE

(A) ORGANISM: Agrobacterium radiobacter

(xi) SEQUENCE DESCRIPTION:

| | | |
|----|---|-----|
| | ATG GAT ATC ATC ATC AAG AAC GGA ACC ATC GTA ACC GCG GAC GGG | 45 |
| | Met Asp Ile Ile Ile Lys Asn Gly Thr Ile Val Thr Ala Asp Gly | |
| 15 | 5 10 15 | |
| | ATT TCT CCC GCC GAT CTC GGA ATC AAG GAT GGC AAG ATC GCC CAG | 90 |
| | Ile Ser Pro Ala Asp Leu Gly Ile Lys Asp Gly Lys Ile Ala Gln | |
| 20 | 20 25 30 | |
| | ATC GGC GGA ACG TTC GGC CCG GCC GGC CGG ACA ATC GAC GCC TCC | 135 |
| | Ile Gly Gly Thr Phe Gly Pro Ala Gly Arg Thr Ile Asp Ala Ser | |
| 25 | 35 40 45 | |
| | GGC CGC TAC GTT TTT CCG GGC GGC ATC GAC GTT CAT ACG CAT GTC | 180 |
| | Gly Arg Tyr Val Phe Pro Gly Gly Ile Asp Val His Thr His Val | |
| 30 | 50 55 60 | |
| | GAG ACG GTC AGC TTC AAC ACG CAG TCG GCC GAC ACA TTC GCA ACC | 225 |
| | Glu Thr Val Ser Phe Asn Thr Gln Ser Ala Asp Thr Phe Ala Thr | |
| 35 | 65 70 75 | |
| | GCG ACG GTC GCG GCC GCC TGT GGC GGC ACG ACG ACC ATC GTC GAT | 270 |
| | Ala Thr Val Ala Ala Ala Cys Gly Gly Thr Thr Thr Ile Val Asp | |
| 40 | 80 85 90 | |
| | TTC TGC CAG CAG GAC CGC GGC CAT AGC CTG AGG GAG GCG GTC GCC | 315 |
| | Phe Cys Gln Gln Asp Arg Gly Hys Ser Leu Arg Glu Ala Val Ala | |
| 45 | 95 100 105 | |
| | AAA TGG GAC GGC ATG GCC GGC GGC AAG TCG GCG ATC GAC TAC GGC | 360 |
| | Lys Trp Asp Gly Met Ala Gly Gly Lys Ser Ala Ile Asp Tyr Gly | |
| 50 | 110 115 120 | |

55

EP 0 677 585 A1

| | | |
|----|---|-----|
| | TAC CAT ATC ATC GTG CTC GAT CCG ACT GAT AGC GTG ATC GAG GAG | 405 |
| | Tyr His Ile Ile Val Leu Asp Pro Thr Asp Ser Val Ile Glu Glu | |
| 5 | 125 130 135 | |
| | CTA GAG GTA CTG CCA GAT CTC GGC ATC ACC TCC TTC AAG GTC TTC | 450 |
| | Leu Glu Val Leu Pro Asp Leu Gly Ile Thr Ser Phe Lys Val Phe | |
| 10 | 140 145 150 | |
| | ATG GCT TAT CGC GGC ATG AAC ATG ATC GAC GAC GTG ACA CTG CTC | 495 |
| | Met Ala Tyr Arg Gly Met Asn Met Ile Asp Asp Val Thr Leu Leu | |
| 15 | 155 160 165 | |
| | AGG ACG CTC GAC AAG GCC GCC AAG ACT GGG TCA CTC GTC ATG GTG | 540 |
| | Arg Thr Leu Asp Lys Ala Ala Lys Thr Gly Ser Leu Val Met Val | |
| 20 | 170 175 180 | |
| | CAC GCG GAG AAC GGC GAC GCC GCC GAC TAT CTT CGC GAC AAG TTC | 585 |
| | His Ala Glu Asn Gly Asp Ala Ala Asp Tyr Leu Arg Asp Lys Phe | |
| 25 | 185 190 195 | |
| | GTC GCC GAT GGC AAA ACG GCG CCG ATC TAC CAC GCG CTC AGC CGT | 630 |
| | Val Ala Asp Gly Lys Thr Ala Pro Ile Tyr His Ala Leu Ser Arg | |
| 30 | 200 205 210 | |
| | CCG CCC CGG GTC GAA GCC GAG GCG ACC GCG AGG GCC CTC GCC CTG | 675 |
| | Pro Pro Arg Val Glu Ala Glu Ala Thr Ala Arg Ala Leu Ala Leu | |
| 35 | 215 220 225 | |
| | GCG GAA ATC GTC AAC GCC CCG ATC TAC ATC GTG CAT CTG ACC TGC | 720 |
| | Ala Glu Ile Val Asn Ala Pro Ile Tyr Ile Val His Leu Thr Cys | |
| 40 | 230 235 240 | |
| | GAA GAA TCC TTC GAC GAG TTG ATG CGG GCA AAG GCT CGG GGT GTC | 765 |
| | Glu Glu Ser Phe Asp Glu Leu Met Arg Ala Lys Ala Arg Gly Val | |
| 45 | 245 250 255 | |
| | CAC GCC CTG GCC GAA ACC TGC ACA CAA TAC CTC TAC CTC ACC AAG | 810 |
| | His Ala Leu Ala Glu Thr Cys Thr Gln Tyr Leu Tyr Leu Thr Lys | |
| 50 | 260 265 270 | |
| 55 | | |

| | | |
|----|---|------|
| | GAC GAC CTG GAG CGG CCG GAT TTC GAG GGC GCG AAG TAT GTT TTC | 855 |
| | Asp Asp Leu Glu Arg Pro Asp Phe Glu Gly Ala Lys Tyr Val Phe | |
| 5 | 275 280 285 | |
| | ACC CCG CCT CCG CGC ACG AAG AAG GAC CAG GAA ATC CTC TGG AAC | 900 |
| | Thr Pro Pro Pro Arg Thr Lys Lys Asp Gln Glu Ile Leu Trp Asn | |
| 10 | 290 295 300 | |
| | GCA CTC CGG AAC GGG GTC CTC GAA ACG GTC TCC TCG GAC CAT TGT | 945 |
| | Ala Leu Arg Asn Gly Val Leu Glu Thr Val Ser Ser Asp His Cys | |
| 15 | 305 310 315 | |
| | TCC TGG CTC TTC GAG GGG CAC AAG GAT CGG GGC AGG AAC GAC TTC | 990 |
| | Ser Trp Leu Phe Glu Gly His Lys Asp Arg Gly Arg Asn Asp Phe | |
| 20 | 320 325 330 | |
| | CGC GCC ATC CCG AAC GGA GCG CCG GGC GTC GAG GAG CGG CTG ATG | 1035 |
| | Arg Ala Ile Pro Asn Gly Ala Pro Gly Val Glu Glu Arg Leu Met | |
| 25 | 335 340 345 | |
| | ATG GTC TAT CAG GGC GTC AAC GAA GGC CGC ATT TCC CTC ACC CAG | 1080 |
| | Met Val Tyr Gln Gly Val Asn Glu Gly Arg Ile Ser Leu Thr Gln | |
| 30 | 350 355 360 | |
| | TTC GTA GAA CTG GTC GCC ACG CGC CCG GCC AAG GTC TTC GGC ATG | 1125 |
| | Phe Val Glu Leu Val Ala Thr Arg Pro Ala Lys Val Phe Gly Met | |
| 35 | 365 370 375 | |
| | TTC CCG GAA AAA GGA ACG GTC CCG GTC GGT TCG GAT GCC GAC ATC | 1170 |
| | Phe Pro Glu Lys Gly Thr Val Ala Val Gly Ser Asp Ala Asp Ile | |
| 40 | 380 385 390 | |
| | GTC CTT TGG GAT CCC GAG GCT GAA ATG GTG ATC GAA CAA AGC GCC | 1215 |
| | Val Leu Trp Asp Pro Glu Ala Glu Met Val Ile Glu Gln Ser Ala | |
| 45 | 395 400 405 | |
| | ATG CAT AAC GCC ATG GAT TAC TCC TCC TAC GAG GGA CAC AAG ATC | 1260 |
| | Met His Asn Ala Met Asp Tyr Ser Ser Tyr Glu Gly His Lys Ile | |
| 50 | 410 415 420 | |

55

EP 0 677 585 A1

AAG GGC GTG CCG AAG ACA GTG CTG CTG CGT GGC AAG GTG ATC GTC 1305

Lys Gly Val Pro Lys Thr Val Leu Leu Arg Gly Lys Val Ile Val

5 425 430 435

GAC GAG GGA ACC TAC GTG GGG GCG CCG ACG GAT GGC CAG TTC CGG 1350

Asp Glu Gly Thr Tyr Val Gly Ala Pro Thr Asp Gly Gln Phe Arg

10 440 445 450

AAG CGC CGC AAA TAC AAG CAA TAA 1373

Lys Arg Arg Lys Tyr Lys Gln STOP

15 455

20

25

30

35

40

45

50

55

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

(A) NAME: ENIRICERCHE S.p.A.
(B) STREET: VIA MARITANO, 26
(C) CITY: SAN DONATO MILANESE, (MILANO)
(E) COUNTRY: ITALY
(F) POSTAL CODE: 20097

TELEFAX: 02/52036344

(ii) TITLE INVENTION:

PROCESS FOR THE PRODUCTION OF D-N-ALFA AMINO ACIDS

(iii) NUMBER OF SEQUENCES: 19

(iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette

(iv) COMPUTER: IBM PC-Compatible

OPERATING SYSTEM: IBM-DOS 5.2/WINDOWS 3.1

SOFTWARE:DisplayWrite 4

(1) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGHT: 23 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA (Genomic)

(xi) SEQUENCE DESCRIPTION:

CGAATTGTAA ATTATGCAGC AGC

23

(1) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGHT: 23 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA (Genomic)

(xi) SEQUENCE DESCRIPTION:

AGGATCGTGA ACTACGCGGC GGC

23

(1) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

5 (A) LENGHT: 23 base pairs
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear
 (ii) MOLECULE TYPE: DNA (Genomic)
 (xi) SEQUENCE DESCRIPTION:
 10 CGCATAGTCA ATTATGCCGC CGC 23
 (1) INFORMATION FOR SEQ ID NO:4:
 (i) SEQUENCE CHARACTERISTICS:
 15 (A) LENGHT: 23 base pairs
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear
 20 (ii) MOLECULE TYPE: DNA (Genomic)
 (xi) SEQUENCE DESCRIPTION:
 CGTATTGTTA ATTATGCTGC TGC 23
 (1) INFORMATION FOR SEQ ID NO:5:
 25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGHT: 17 base pairs
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
 30 (D) TOPOLOGY: Linear
 (ii) MOLECULE TYPE: DNA (Genomic)
 (xi) SEQUENCE DESCRIPTION:
 35 GGACCAATTC AACGAGC 17
 (1) INFORMATION FOR SEQ ID NO:6:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGHT: 17 base pairs
 40 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear
 (ii) MOLECULE TYPE: DNA (Genomic)
 45 (xi) SEQUENCE DESCRIPTION:
 GGGCCGATCC AGCGGGC 17
 (1) INFORMATION FOR SEQ ID NO:7:
 (i) SEQUENCE CHARACTERISTICS:
 50 (A) LENGHT: 17 base pairs

55

5 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear
 (ii) MOLECULE TYPE: DNA (Genomic)
 (xi) SEQUENCE DESCRIPTION:
 GGCCCCATAC AACGCGC 17

10 (1) INFORMATION FOR SEQ ID NO:8:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGHT: 17 base pairs
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear
 (ii) MOLECULE TYPE: DNA (Genomic)
 (xi) SEQUENCE DESCRIPTION:
 20 GGTCTTATTC AACGTGC 17
 (1) INFORMATION FOR SEQ ID NO:9:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGHT: 17 base pairs
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear
 (ii) MOLECULE TYPE: DNA (Genomic)
 (xi) SEQUENCE DESCRIPTION:
 30 CGAGCAGATG TAATGGA 17
 (1) INFORMATION FOR SEQ ID NO:10:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGHT: 17 base pairs
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear
 (ii) MOLECULE TYPE: DNA (Genomic)
 (xi) SEQUENCE DESCRIPTION:
 40 AGGGCGGACG TGATGGA 17
 (1) INFORMATION FOR SEQ ID NO:11:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGHT: 17 base pairs
 (B) TYPE: Nucleic acid

55

(C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear
 (ii) MOLECULE TYPE: DNA (Genomic)
 5 (xi) SEQUENCE DESCRIPTION:
 CGCGCCGATG TCATGGA 17
 (1) INFORMATION FOR SEQ ID NO:12:
 10 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGHT: 17 base pairs
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
 15 (D) TOPOLOGY: Linear
 (ii) MOLECULE TYPE: DNA (Genomic)
 (xi) SEQUENCE DESCRIPTION:
 CGTGCTGATG TTATGGA 17
 20 (1) INFORMATION FOR SEQ ID NO:13:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGHT: 27 base pairs
 (B) TYPE: Nucleic acid
 25 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear
 (ii) MOLECULE TYPE: DNA (Genomic)
 30 (xi) SEQUENCE DESCRIPTION:
 ATCGTAACCG CGGACGGGAT TTCTCCC 27
 (1) INFORMATION FOR SEQ ID NO:14:
 (i) SEQUENCE CHARACTERISTICS:
 35 (A) LENGHT: 22 base pairs
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear
 40 (ii) MOLECULE TYPE: DNA (Genomic)
 (ix) FEATURE:
 (A) NAME: Primer
 45 (xi) SEQUENCE DESCRIPTION:
 GGGAATTCTT ATGACACGTC AG 22
 (1) INFORMATION FOR SEQ ID NO:15:
 (i) SEQUENCE CHARACTERISTICS:
 50 (A) LENGHT: 23 base pairs

55

(B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
 5 (D) TOPOLOGY: Linear
 (ii) MOLECULE TYPE: DNA (Genomic)
 (ix) FEATURE:
 (A) NAME: Primer
 10 (xi) SEQUENCE DESCRIPTION:
 CCCAAGCTTC AAAATTCCGC GAT 23
 (1) INFORMATION FOR SEQ ID NO:16:
 15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGHT: 13 base pairs
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
 20 (D) TOPOLOGY: Linear
 (ii) MOLECULE TYPE: DNA (Genomic)
 (ix) FEATURE:
 (A) NAME: Linker
 25 (xi) SEQUENCE DESCRIPTION:
 AATTCTTATG GAT 13
 (1) INFORMATION FOR SEQ ID NO:17:
 30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGHT: 76 base pairs
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
 35 (D) TOPOLOGY: Linear
 (ii) MOLECULE TYPE: DNA (Genomic)
 (ix) FEATURE:
 (A) NAME: Linker
 40 (xi) SEQUENCE DESCRIPTION:
 TCGACGAGGG AACCTACGTG GGGGCGCCGA CGGATGGCCA 40
 GTTCCGGAAG CGCCGCAAT ACAAGCAATA AGGAGG 76
 45 (1) INFORMATION FOR SEQ ID NO:18:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGHT: 915 base pairs
 (B) TYPE: Nucleic acid
 50 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear

55

(ii) MOLECULE TYPE: DNA (Genomic)

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Agrobacterium radiobacter

(xi) SEQUENCE DESCRIPTION:

| | | |
|----|---|-----|
| | ATG ACA CGT CAG ATG ATA CTT GCT GTC GGA CAG CAA GGC CCC ATC | 45 |
| | Met Thr Arg Gln Met Ile Leu Ala Val Gly Gln Gln Gly Pro Ile | |
| 10 | 5 10 15 | |
| | GCG CGA GCG GAG ACA CGC GAA CAG GTG GTT GGC CGC CTC CTC GAC | 90 |
| | Ala Arg Ala Glu Thr Arg Glu Gln Val Val Gly Arg Leu Leu Asp | |
| | 20 25 30 | |
| 15 | ATG TTG ACG AAC GCA GCC AGC CGG GGC GTG AAC TTC ATC GTC TTT | 135 |
| | Met Leu Thr Asn Ala Ala Ser Arg Gly Val Asn Phe Ile Val Phe | |
| | 35 40 45 | |
| 20 | CCC GAG CTT GCG CTC ACG ACC TTC TTC CCG CGC TGG CAT TTC ACC | 180 |
| | Pro Glu Leu Ala Leu Thr Thr Phe Phe Pro Arg Trp His Phe Thr | |
| | 50 55 60 | |
| | GAC GAG GCC GAG CTC GAT AGC TTC TAT GAG ACC GAA ATG CCC GGC | 225 |
| 25 | Asp Glu Ala Glu Leu Asp Ser Phe Tyr Glu Thr Glu Met Pro Gly | |
| | 65 70 75 | |
| | CCG GTG GTC CGT CCA CTC TTT GAG ACG GCC GCC GAA CTC GGG ATC | 270 |
| | Pro Val Val Arg Pro Leu Phe Glu Thr Ala Ala Glu Leu Gly Ile | |
| 30 | 80 85 90 | |
| | GGC TTC AAT CTG GGC TAC GCC GAA CTC GTC GTC GAA GGC GGC GTC | 315 |
| | Gly Phe Asn Leu Gly Tyr Ala Glu Leu Val Val Glu Gly Gly Val | |
| | 95 100 105 | |
| 35 | AAG CGT CGC TTC AAC ACG TCC ATT CTG GTG GAT AAG TCA GGC AAG | 360 |
| | Lys Arg Arg Phe Asn Thr Ser Ile Leu Val Asp Lys Ser Gly Lys | |
| | 110 115 120 | |
| 40 | ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG GGT CAC AAG GAG | 405 |
| | Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro Gly Hys Lys Glu | |
| | 125 130 135 | |
| | TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG CGT TAT TTC | 450 |
| 45 | Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys Arg Tyr Phe | |
| | 140 145 150 | |
| | GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC GCC GCG | 495 |
| 50 | Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp Ala Ala | |
| | 155 160 165 | |

55

EP 0 677 585 A1

AAA ATG GGG ATG TTC ATC TGC AAC GAT CGC CGC TGG CCT GAA ACG 540
 Lys Met Gly Met Phe Ile Cys Asn Asp Arg Arg Trp Pro Glu Thr
 170 175 180
 5 TGG CGG GTG ATG GGA CTT AAG GGC GCC GAG ATC ATC TGC GGC GGC 585
 Trp Arg Val Met Gly Leu Lys Gly Ala Glu Ile Ile Cys Gly Gly
 185 190 195
 10 TAC AAC ACG CCG ACC CAC AAT CCC CCC GTT CCC CAG CAC GAC CAT 630
 Tyr Asn Thr Pro Thr His Asn Pro Pro Val Pro Gln His Asp His
 200 205 210
 CTG ACG TCC TTC CAC CAC CTT CTG TCG ATG CAG GCC GGG TCG TAC 675
 15 Leu Thr Ser Phe His His Leu Leu Ser Met Gln Ala Gly Ser Tyr
 215 220 225
 CAA AAC GGC GCC TGG TCC GCG GCG GCC GGC AAG GTC GGC ATG GAG 720
 Gln Asn Gly Ala Trp Ser Ala Ala Ala Gly Lys Val Gly Met Glu
 20 230 235 240
 GAG GGG TGC ATG CTG CTC GGC CAT TCG TGC ATC GTG GCG CCG ACC 765
 Glu Gly Cys Met Leu Leu Gly His Ser Cys Ile Val Ala Pro Thr
 245 250 255
 25 GGC GAA ATC GTT GCC CTG ACC ACG ACG TTG GAA GAC GAG GTG ATC 810
 Gly Glu Ile Val Ala Leu Thr Thr Thr Leu Glu Asp Glu Val Ile
 260 265 270
 30 ACC GCC GCC GTC GAT CTC GAC CGC TGC CGG GAA CTG CGC GAA CAC 855
 Thr Ala Ala Val Asp Leu Asp Arg Cys Arg Glu Leu Arg Glu His
 275 280 285
 ATC TTC AAT TTC AAA GCC CAT CGT CAG CCA CAG CAC TAC GGT CTG 900
 35 Ile Phe Asn Phe Lys Ala His Arg Gln Pro Gln His Tyr Gly Leu
 290 295 300
 ATC GCG GAA TTT TGA 915
 40 Ile Ala Glu Phe
 (1) INFORMATION FOR SEQ ID NO:19:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGHT: 1374 base pairs
 45 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear
 (ii) MOLECULE TYPE: DNA (Genomic)
 50 (vi) ORIGINAL SOURCE

55

(A) ORGANISM: Agrobacterium radiobacter

(xi) SEQUENCE DESCRIPTION:

```

5  ATG GAT ATC ATC ATC AAG AAC GGA ACC ATC GTA ACC GCG GAC GGG 45
   Met Asp Ile Ile Ile Lys Asn Gly Thr Ile Val Thr Ala Asp Gly
      5              10              15
10  ATT TCT CCC GCC GAT CTC GGA ATC AAG GAT GGC AAG ATC GCC CAG 90
   Ile Ser Pro Ala Asp Leu Gly Ile Lys Asp Gly Lys Ile Ala Gln
      20              25              30
15  ATC GGC GGA ACG TTC GGC CCG GCC GGC CGG ACA ATC GAC GCC TCC 135
   Ile Gly Gly Thr Phe Gly Pro Ala Gly Arg Thr Ile Asp Ala Ser
      35              40              45
20  GGC CGC TAC GTT TTT CCG GGC GGC ATC GAC GTT CAT ACG CAT GTC 180
   Gly Arg Tyr Val Phe Pro Gly Gly Ile Asp Val His Thr His Val
      50              55              60
25  GAG ACG GTC AGC TTC AAC ACG CAG TCG GCC GAC ACA TTC GCA ACC 225
   Glu Thr Val Ser Phe Asn Thr Gln Ser Ala Asp Thr Phe Ala Thr
      65              70              75
30  GCG ACG GTC GCG GCC GCC TGT GGC GGC ACG ACG ACC ATC GTC GAT 270
   Ala Thr Val Ala Ala Ala Cys Gly Gly Thr Thr Thr Ile Val Asp
      80              85              90
35  TTC TGC CAG CAG GAC CGC GGC CAT AGC CTG AGG GAG GCG GTC GCC 315
   Phe Cys Gln Gln Asp Arg Gly Hys Ser Leu Arg Glu Ala Val Ala
      95              100             105
40  AAA TGG GAC GGC ATG GCC GGC GGC AAG TCG GCG ATC GAC TAC GGC 360
   Lys Trp Asp Gly Met Ala Gly Gly Lys Ser Ala Ile Asp Tyr Gly
      110             115             120
45  TAC CAT ATC ATC GTG CTC GAT CCG ACT GAT AGC GTG ATC GAG GAG 405
   Tyr His Ile Ile Val Leu Asp Pro Thr Asp Ser Val Ile Glu Glu
      125             130             135
50  CTA GAG GTA CTG CCA GAT CTC GGC ATC ACC TCC TTC AAG GTC TTC 450
   Leu Glu Val Leu Pro Asp Leu Gly Ile Thr Ser Phe Lys Val Phe
      140             145             150
55  ATG GCT TAT CGC GGC ATG AAC ATG ATC GAC GAC GTG ACA CTG CTC 495
   Met Ala Tyr Arg Gly Met Asn Met Ile Asp Asp Val Thr Leu Leu
      155             160             165
60  AGG ACG CTC GAC AAG GCC GCC AAG ACT GGG TCA CTC GTC ATG GTG 540
   Arg Thr Leu Asp Lys Ala Ala Lys Thr Gly Ser Leu Val Met Val

```

EP 0 677 585 A1

| | | | | | | | |
|----|-----|-----|-----|-----|-----|-----|--|
| | | 170 | | 175 | | 180 | |
| | CAC | GCG | GAG | AAC | GGC | GAC | GCC GCC GAC TAT CTT CGC GAC AAG TTC 585 |
| 5 | His | Ala | Glu | Asn | Gly | Asp | Ala Ala Asp Tyr Leu Arg Asp Lys Phe |
| | | 185 | | 190 | | 195 | |
| | GTC | GCC | GAT | GGC | AAA | ACG | GCG CCG ATC TAC CAC GCG CTC AGC CGT 630 |
| 10 | Val | Ala | Asp | Gly | Lys | Thr | Ala Pro Ile Tyr His Ala Leu Ser Arg |
| | | 200 | | 205 | | 210 | |
| | CCG | CCC | CGG | GTC | GAA | GCC | GAG GCG ACC GCG AGG GCC CTC GCC CTG 675 |
| | Pro | Pro | Arg | Val | Glu | Ala | Glu Ala Thr Ala Arg Ala Leu Ala Leu |
| 15 | | 215 | | 220 | | 225 | |
| | GCG | GAA | ATC | GTC | AAC | GCC | CCG ATC TAC ATC GTG CAT CTG ACC TGC 720 |
| | Ala | Glu | Ile | Val | Asn | Ala | Pro Ile Tyr Ile Val His Leu Thr Cys |
| | | 230 | | 235 | | 240 | |
| 20 | GAA | GAA | TCC | TTC | GAC | GAG | TTG ATG CGG GCA AAG GCT CGG GGT GTC 765 |
| | Glu | Glu | Ser | Phe | Asp | Glu | Leu Met Arg Ala Lys Ala Arg Gly Val |
| | | 245 | | 250 | | 255 | |
| | CAC | GCC | CTG | GCC | GAA | ACC | TGC ACA CAA TAC CTC TAC CTC ACC AAG 810 |
| 25 | His | Ala | Leu | Ala | Glu | Thr | Cys Thr Gln Tyr Leu Tyr Leu Thr Lys |
| | | 260 | | 265 | | 270 | |
| | GAC | GAC | CTG | GAG | CGG | CCG | GAT TTC GAG GGC GCG AAG TAT GTT TTC 855 |
| 30 | Asp | Asp | Leu | Glu | Arg | Pro | Asp Phe Glu Gly Ala Lys Tyr Val Phe |
| | | 275 | | 280 | | 285 | |
| | ACC | CCG | CCT | CCG | CGC | ACG | AAG AAG GAC CAG GAA ATC CTC TGG AAC 900 |
| | Thr | Pro | Pro | Pro | Arg | Thr | Lys Lys Asp Gln Glu Ile Leu Trp Asn |
| 35 | | 290 | | 295 | | 300 | |
| | GCA | CTC | CGG | AAC | GGG | GTC | CTC GAA ACG GTC TCC TCG GAC CAT TGT 945 |
| | Ala | Leu | Arg | Asn | Gly | Val | Leu Glu Thr Val Ser Ser Asp His Cys |
| | | 305 | | 310 | | 315 | |
| 40 | TCC | TGG | CTC | TTC | GAG | GGG | CAC AAG GAT CGG GGC AGG AAC GAC TTC 990 |
| | Ser | Trp | Leu | Phe | Glu | Gly | His Lys Asp Arg Gly Arg Asn Asp Phe |
| | | 320 | | 325 | | 330 | |
| | CGC | GCC | ATC | CCG | AAC | GGA | GCG CCG GGC GTC GAG GAG CGG CTG ATG 1035 |
| 45 | Arg | Ala | Ile | Pro | Asn | Gly | Ala Pro Gly Val Glu Glu Arg Leu Met |
| | | 335 | | 340 | | 345 | |
| | ATG | GTC | TAT | CAG | GGC | GTC | AAC GAA GGC CGC ATT TCC CTC ACC CAG 1080 |
| 50 | Met | Val | Tyr | Gln | Gly | Val | Asn Glu Gly Arg Ile Ser Leu Thr Gln |
| | | 350 | | 355 | | 360 | |

55

TTC GTA GAA CTG GTC GCC ACG CGC CCG GCC AAG GTC TTC GGC ATG 1125
 Phe Val Glu Leu Val Ala Thr Arg Pro Ala Lys Val Phe Gly Met
 5 365 370 375
 TTC CCG GAA AAA GGA ACG GTC GCG GTC GGT TCG GAT GCC GAC ATC 1170
 Phe Pro Glu Lys Gly Thr Val Ala Val Gly Ser Asp Ala Asp Ile
 380 385 390
 10 GTC CTT TGG GAT CCC GAG GCT GAA ATG GTG ATC GAA CAA AGC GCC 1215
 Val Leu Trp Asp Pro Glu Ala Glu Met Val Ile Glu Gln Ser Ala
 395 400 405
 15 ATG CAT AAC GCC ATG GAT TAC TCC TCC TAC GAG GGA CAC AAG ATC 1260
 Met His Asn Ala Met Asp Tyr Ser Ser Tyr Glu Gly His Lys Ile
 410 415 420
 AAG GGC GTG CCG AAG ACA GTG CTG CTG CGT GGC AAG GTG ATC GTC 1305
 20 Lys Gly Val Pro Lys Thr Val Leu Leu Arg Gly Lys Val Ile Val
 425 430 435
 GAC GAG GGA ACC TAC GTG GGG GCG CCG ACG GAT GGC CAG TTC CGG 1350
 Asp Glu Gly Thr Tyr Val Gly Ala Pro Thr Asp Gly Gln Phe Arg
 25 440 445 450
 AAG CGC CGC AAA TAC AAG CAA TAA 1374
 Lys Arg Arg Lys Tyr Lys Gln
 30 455

Claims

- 35
1. A process for the production of D- α -amino acids by the stereospecific conversion of racemic mixtures of 5-substituted hydantoins characterized in that, the conversion reaction is carried out in the presence of a microorganism transformed with the plasmid pSM651 CBS 203.94 capable of expressing at high levels and without inducers an enzymatic system capable of converting said hydantoins into the corresponding D- α -amino acids.
 - 40 2. The process according to claim 1, characterized in that, the conversion reaction is carried out in the presence of the enzymatic system isolated from a microorganism transformed with the plasmid pSM651 CBS 203.94.
 - 45 3. The process according to claim 2, characterized in that, said enzymatic system is immobilized on an insoluble support.
 - 50 4. The process according to claim 1, characterized in that, the microorganisms are selected from the group of Bacillus subtilis and Escherichia coli.
 - 55 5. The process according to claim 1, characterized in that, the 5-substituted hydantoin is selected from D,L-5-phenylhydantoin, D,L-5-para-hydroxyphenylhydantoin, D,L-5-methylhydantoin, D,L-5-isopropylhydantoin, D,L-5-thienylhydantoin, D,L-5-para-methoxyphenylhydantoin, D,L-5-para-chloro phenylhydantoin, D,L-5-benzylhydantoin.
 6. The process according to claim 5, characterized in that, the hydantoin is D,L-5-para-hydroxyphenylhydantoin.

7. The process according to claim 5, characterized in that, the hydantoin is D,L-5-phenylhydantoin.
8. The process according to claim 1, characterized in that, the conversion reaction is carried out at a temperature of between 20 °C and 60 °C.
9. The process according to claim 8, characterized in that, the temperature is between 30 ° and 45 °C.
10. The process according to claim 1, characterized in that, the conversion reaction is carried out at a pH of between 6.0 and 10.
11. The process according to claim 10, characterized in that, the pH is between 7.0 and 8.5.
12. The process according to claim 1, characterized in that, the conversion reaction is carried out using a weight ratio biomass/hydantoins of between 1/1 and 1/50.
13. Plasmid pSM651 deposited at the Bureau Voor Schimmelcultures, SK Baarn (Holland) where it has received the deposit number CBS 203.94.
14. A microorganism selected from Bacillus subtilis and Escherichia coli transformed with the plasmid pSM651.
15. The microorganism according to claim 10, which is Escherichia coli SMC305 CBS 203.94.

FIG. 1

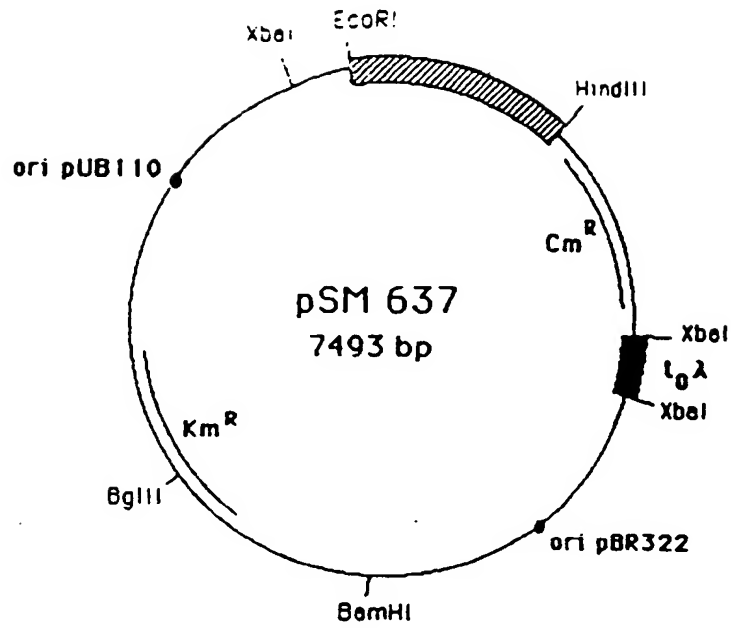


FIG. 2

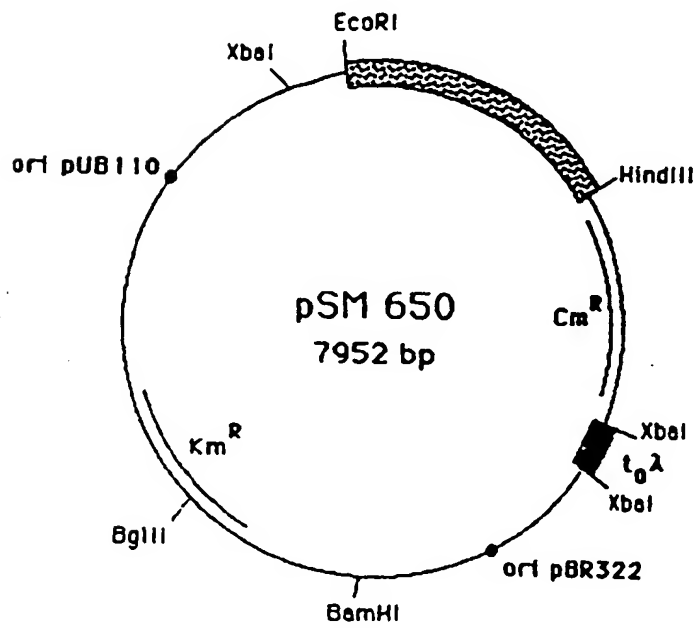


FIG. 3

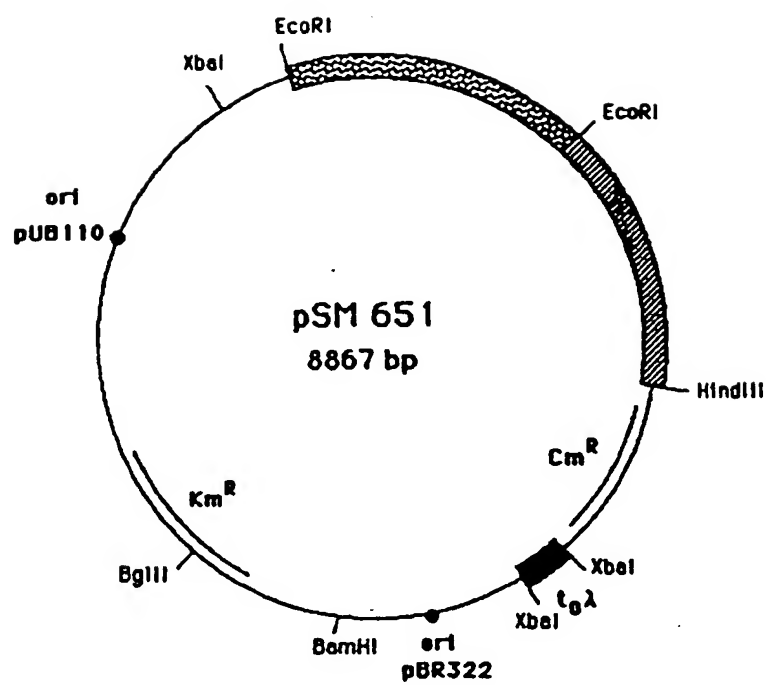


FIG. 4A

| | |
|---|-----|
| ATG ACA CGT CAG ATG ATA CTT GCT GTC GGA CAG CAA GGC CCC ATC | 45 |
| Met Thr Arg Gln Met Ile Leu Ala Val Gly Gln Gln Gly Pro Ile | |
| 5 10 15 | |
| GCG CGA GCG GAG ACA CGC GAA CAG GTG GTT GGC CGC CTC CTC GAC | 90 |
| Ala Arg Ala Glu Thr Arg Glu Gln Val Val Gly Arg Leu Leu Asp | |
| 20 25 30 | |
| ATG TTG ACG AAC GCA GCC AGC CGG GGC GTG AAC TTC ATC GTC TTT | 135 |
| Met Leu Thr Asn Ala Ala Ser Arg Gly Val Asn Phe Ile Val Phe | |
| 35 40 45 | |
| CCC GAG CTT GCG CTC ACG ACC TTC TTC CCG CGC TGG CAT TTC ACC | 180 |
| Pro Glu Leu Ala Leu Thr Thr Phe Phe Pro Arg Trp His Phe Thr | |
| 50 55 60 | |
| GAC GAG GCC GAG CTC GAT AGC TTC TAT GAG ACC GAA ATG CCC GGC | 225 |
| Asp Glu Ala Glu Leu Asp Ser Phe Tyr Glu Thr Glu Met Pro Gly | |
| 65 70 75 | |
| CCG GTG GTC CGT CCA CTC TTT GAG ACG GCC GCC GAA CTC GGG ATC | 270 |
| Pro Val Val Arg Pro Leu Phe Glu Thr Ala Ala Glu Leu Gly Ile | |
| 80 85 90 | |
| GGC TTC AAT CTG GGC TAC GCC GAA CTC GTC GTC GAA GGC GGC GTC | 315 |
| Gly Phe Asn Leu Gly Tyr Ala Glu Leu Val Val Glu Gly Gly Val | |
| 95 100 105 | |
| AAG CGT CGC TTC AAC ACG TCC ATT CTG GTG GAT AAG TCA GGC AAG | 360 |
| Lys Arg Arg Phe Asn Thr Ser Ile Leu Val Asp Lys Ser Gly Lys | |
| 110 115 120 | |
| ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG GGT CAC AAG GAG | 415 |
| Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro Gly His Lys Glu | |
| 125 130 135 | |
| TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG CGT TAT TTC | 450 |
| Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys Arg Tyr Phe | |
| 140 145 150 | |

FIG. 4B

| | |
|---|-----|
| GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC GCC GCG | 495 |
| Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp Ala Ala | |
| 155 160 165 | |
| AAA ATG GGG ATG TTC ATC TGC AAC GAT CGC CGC TGG CCT GAA ACG | 540 |
| Lys Met Gly Met Phe Ile Cys Asn Asp Arg Arg Trp Pro Glu Thr | |
| 170 175 180 | |
| TGG CGG GTG ATG GGA CTT AAG GGC GCC GAG ATC ATC TGC GGC GGC | 585 |
| Trp Arg Val Met Gly Leu Lys Gly Ala Glu Ile Ile Cys Gly Gly | |
| 185 190 195 | |
| TAC AAC ACG CCG ACC CAC AAT CCC CCC GTT CCC CAG CAC GAC CAT | 630 |
| Tyr Asn Thr Pro Thr His Asn Pro Pro Val Pro Gln His Asp His | |
| 200 205 210 | |
| CTG ACG TCC TTC CAC CAC CTT CTG TCG ATG CAG GCC GGG TCG TAC | 675 |
| Leu Thr Ser Phe His His Leu Leu Ser Met Gln Ala Gly Ser Tyr | |
| 215 220 225 | |
| CAA AAC GGC GCC TGG TCC GCG GCG GCC GGC AAG GTC GGC ATG GAG | 720 |
| Gln Asn Gly Ala Trp Ser Ala Ala Ala Gly Lys Val Gly Met Glu | |
| 230 235 240 | |
| GAG GGC TGC ATG CTG CTC GGC CAT TCG TGC ATC GTG GCG CCG ACC | 765 |
| Glu Gly Cys Met Leu Leu Gly His Ser Cys Ile Val Ala Pro Thr | |
| 245 250 255 | |
| GGC GAA ATC GTT GCC CTG ACC ACG ACG TTG GAA GAC GAG GTG ATC | 810 |
| Gly Glu Ile Val Ala Leu Thr Thr Thr Leu Glu Asp Glu Val Ile | |
| 260 265 270 | |
| ACC GCC GCC GTC GAT CTC GAC CGC TGC CGG GAA CTG CGC GAA CAC | 855 |
| Thr Ala Ala Val Asp Leu Asp Arg Cys Arg Glu Leu Arg Glu His | |
| 275 280 285 | |
| ATC TTC AAT TTC AAA GCC CAT CGT CAG CCA CAG CAC TAC GGT CTG | 900 |
| Ile Phe Asn Phe Lys Ala His Arg Gln Pro Gln His Tyr Gly Leu | |
| 290 295 300 | |
| ATC GCG GAA TTT TGA | 915 |
| Ile Ala Glu Phe STOP | |

FIG. 5A

| | |
|---|-----|
| ATG GAT ATC ATC ATC AAG AAC GGA ACC ATC GTA ACC GCG GAC GGG | 45 |
| Met Asp Ile Ile Ile Lys Asn Gly Thr Ile Val Thr Ala Asp Gly | |
| 5 10 15 | |
| ATT TCT CCC GCC GAT CTC GGA ATC AAG GAT GGC AAG ATC GCC CAG | 90 |
| Ile Ser Pro Ala Asp Leu Gly Ile Lys Asp Gly Lys Ile Ala Gln | |
| 20 25 30 | |
| ATC GGC GGA ACG TTC GGC CCG GCC GGC CGG ACA ATC GAC GCC TCC | 135 |
| Ile Gly Gly Thr Phe Gly Pro Ala Gly Arg Thr Ile Asp Ala Ser | |
| 35 40 45 | |
| GGC CGC TAC GTT TTT CCG GGC GGC ATC GAC GTT CAT ACG CAT GTC | 180 |
| Gly Arg Tyr Val Phe Pro Gly Gly Ile Asp Val His Thr His Val | |
| 50 55 60 | |
| GAG ACG GTC AGC TTC AAC ACG CAG TCG GCC GAC ACA TTC GCA ACC | 225 |
| Glu Thr Val Ser Phe Asn Thr Gln Ser Ala Asp Thr Phe Ala Thr | |
| 65 70 75 | |
| GCG ACG GTC GCG GCC GCC TGT GGC GGC ACG ACG ACC ATC GTC GAT | 270 |
| Ala Thr Val Ala Ala Ala Cys Gly Gly Thr Thr Thr Ile Val Asp | |
| 80 85 90 | |
| TTC TGC CAG CAG GAC CGC GGC CAT AGC CTG AGG GAG GCG GTC GCC | 315 |
| Phe Cys Gln Gln Asp Arg Gly Hys Ser Leu Arg Glu Ala Val Ala | |
| 95 100 105 | |
| AAA TGG GAC GGC ATG GCC GGC GGC AAG TCG GCG ATC GAC TAC GGC | 360 |
| Lys Trp Asp Gly Met Ala Gly Gly Lys Ser Ala Ile Asp Tyr Gly | |
| 110 115 120 | |
| TAC CAT ATC ATC GTG CTC GAT CCG ACT GAT AGC GTG ATC GAG GAG | 405 |
| Tyr His Ile Ile Val Leu Asp Pro Thr Asp Ser Val Ile Glu Glu | |
| 125 130 135 | |
| CTA GAG GTA CTG CCA GAT CTC GGC ATC ACC TCC TTC AAG GTC TTC | 450 |
| Leu Glu Val Leu Pro Asp Leu Gly Ile Thr Ser Phe Lys Val Phe | |

FIG. 5B

| | | | | |
|---|-----|-----|-----|-----|
| | 140 | 145 | 150 | |
| ATG GCT TAT CGC GGC ATG AAC ATG ATC GAC GAC GTG ACA CTG CTC | | | | 495 |
| Met Ala Tyr Arg Gly Met Asn Met Ile Asp Asp Val Thr Leu Leu | | | | |
| | 155 | 160 | 165 | |
| AGG ACG CTC GAC AAG GCC GCC AAG ACT GGG TCA CTC GTC ATG GTG | | | | 540 |
| Arg Thr Leu Asp Lys Ala Ala Lys Thr Gly Ser Leu Val Met Val | | | | |
| | 170 | 175 | 180 | |
| CAC GCG GAG AAC GGC GAC GCC GCC GAC TAT CTT CGC GAC AAG TTC | | | | 585 |
| His Ala Glu Asn Gly Asp Ala Ala Asp Tyr Leu Arg Asp Lys Phe | | | | |
| | 185 | 190 | 195 | |
| GTC GCC GAT GGC AAA ACG GCG CCG ATC TAC CAC GCG CTC AGC CGT | | | | 630 |
| Val Ala Asp Gly Lys Thr Ala Pro Ile Tyr His Ala Leu Ser Arg | | | | |
| | 200 | 205 | 210 | |
| CCG CCC CGG GTC GAA GCC GAG GCG ACC GCG AGG GCC CTC GCC CTG | | | | 675 |
| Pro Pro Arg Val Glu Ala Glu Ala Thr Ala Arg Ala Leu Ala Leu | | | | |
| | 215 | 220 | 225 | |
| GCG GAA ATC GTC AAC GCC CCG ATC TAC ATC GTG CAT CTG ACC TGC | | | | 720 |
| Ala Glu Ile Val Asn Ala Pro Ile Tyr Ile Val His Leu Thr Cys | | | | |
| | 230 | 235 | 240 | |
| GAA GAA TCC TTC GAC GAG TTG ATG CCG GCA AAG GCT CGG GGT GTC | | | | 765 |
| Glu Glu Ser Phe Asp Glu Leu Met Arg Ala Lys Ala Arg Gly Val | | | | |
| | 245 | 250 | 255 | |
| CAC GCC CTG GCC GAA ACC TGC ACA CAA TAC CTC TAC CTC ACC AAG | | | | 810 |
| His Ala Leu Ala Glu Thr Cys Thr Gln Tyr Leu Tyr Leu Thr Lys | | | | |
| | 260 | 265 | 270 | |
| GAC GAC CTG GAG CGG CCG GAT TTC GAG GGC GCG AAG TAT GTT TTC | | | | 855 |
| Asp Asp Leu Glu Arg Pro Asp Phe Glu Gly Ala Lys Tyr Val Phe | | | | |
| | 275 | 280 | 285 | |
| ACC CCG CCT CCG CGC ACG AAG AAG GAC CAG GAA ATC CTC TGG AAC | | | | 900 |
| Thr Pro Pro Pro Arg Thr Lys Lys Asp Gln Glu Ile Leu Trp Asn | | | | |
| | 290 | 295 | 300 | |
| GCA CTC CGG AAC GGG GTC CTC GAA ACG GTC TCC TCG GAC CAT TGT | | | | 945 |
| Ala Leu Arg Asn Gly Val Leu Glu Thr Val Ser Ser Asp His Cys | | | | |

FIG. 5C

| | | | |
|---|-----|-----|------|
| 305 | 310 | 315 | |
| TCC TGG CTC TTC GAG GGG CAC AAG GAT CGG GGC AGG AAC GAC TTC | | | 990 |
| Ser Trp Leu Phe Glu Gly His Lys Asp Arg Gly Arg Asn Asp Phe | | | |
| 320 | 325 | 330 | |
| CGC GCC ATC CCG AAC GGA GCG CCG GGC GTC GAG GAG CCG CTG ATG | | | 1035 |
| Arg Ala Ile Pro Asn Gly Ala Pro Gly Val Glu Glu Arg Leu Met | | | |
| 335 | 340 | 345 | |
| ATG GTC TAT CAG GGC GTC AAC GAA GGC CGC ATT TCC CTC ACC CAG | | | 1080 |
| Met Val Tyr Gln Gly Val Asn Glu Gly Arg Ile Ser Leu Thr Gln | | | |
| 350 | 355 | 360 | |
| TTC GTA GAA CTG GTC GCC ACG CGC CCG GCC AAG GTC TTC GGC ATG | | | 1125 |
| Phe Val Glu Leu Val Ala Thr Arg Pro Ala Lys Val Phe Gly Met | | | |
| 365 | 370 | 375 | |
| TTC CCG GAA AAA GGA ACG GTC GCG GTC GGT TCG GAT GCC GAC ATC | | | 1170 |
| Phe Pro Glu Lys Gly Thr Val Ala Val Gly Ser Asp Ala Asp Ile | | | |
| 380 | 385 | 390 | |
| GTC CTT TGG GAT CCC GAG GCT GAA ATG GTG ATC GAA CAA AGC GCC | | | 1215 |
| Val Leu Trp Asp Pro Glu Ala Glu Met Val Ile Glu Gln Ser Ala | | | |
| 395 | 400 | 405 | |
| ATG CAT AAC GCC ATG GAT TAC TCC TCC TAC GAG GGA CAC AAG ATC | | | 1260 |
| Met His Asn Ala Met Asp Tyr Ser Ser Tyr Glu Gly His Lys Ile | | | |
| 410 | 415 | 420 | |
| AAG GGC GTG CCG AAG ACA GTG CTG CTG CGT GGC AAG GTG ATC GTC | | | 1305 |
| Lys Gly Val Pro Lys Thr Val Leu Leu Arg Gly Lys Val Ile Val | | | |
| 425 | 430 | 435 | |
| GAC GAG GGA ACC TAC GTG GGG GCG CCG ACC GAT GGC CAG TTC CCG | | | 1350 |
| Asp Glu Gly Thr Tyr Val Gly Ala Pro Thr Asp Gly Gln Phe Arg | | | |
| 440 | 445 | 450 | |
| AAG CGC CGC AAA TAC AAG CAA TAA | | | 1373 |
| Lys Arg Arg Lys Tyr Lys Gln STOP | | | |
| 455 | | | |

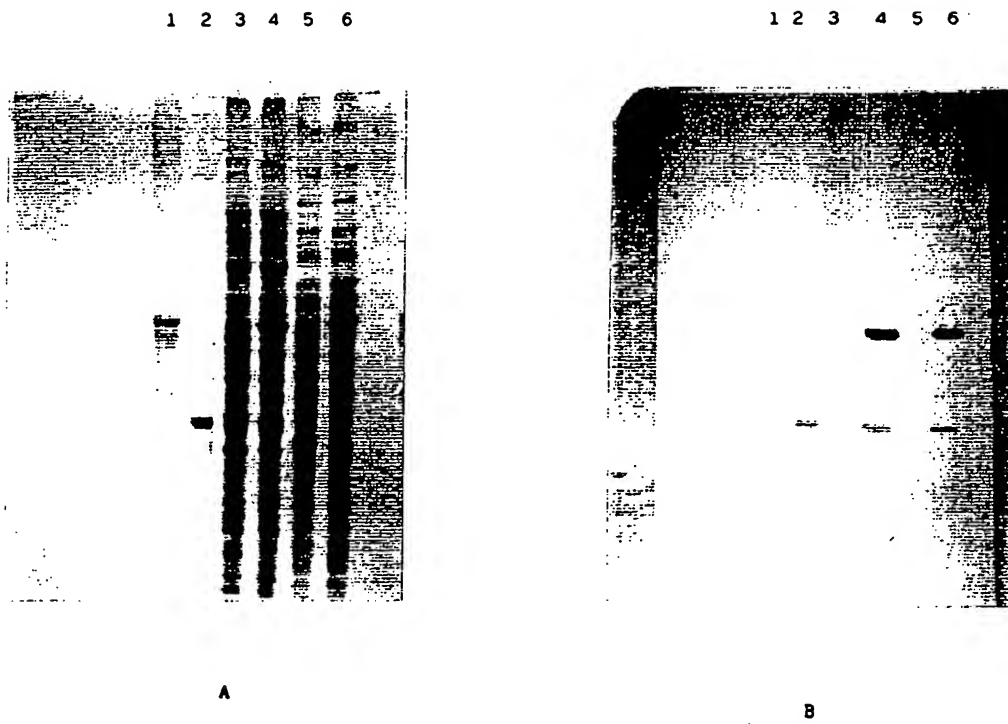


FIG. 6



European Patent
Office

EUROPEAN SEARCH REPORT

Application Number
EP 95 10 4393

| DOCUMENTS CONSIDERED TO BE RELEVANT | | | |
|---|---|--|---|
| Category | Citation of document with indication, where appropriate, of relevant passages | Relevant to claim | CLASSIFICATION OF THE APPLICATION (Int.Cl.6) |
| X | WO-A-94 00577 (SMITHKLINE BEECHAM PLC ; SMITHKLINE BEECHAM CORP (US); NEAL ROBERT) 6 January 1994 * page 1, line 29 - page 3, paragraph 5 * * page 9, line 10 - line 27 * * examples 4, 5, 12, 13, 15, 26-31 * ---- | 1,2,4,8, 9,14 | C12N15/55 C12P41/00 C12N1/21 /(C12N1/21, C12R1:19), (C12P41/00, C12R1:19) |
| D,A | EP-A-0 515 698 (KANEKAFUCHI KAGAKU KOGYO KABUSHIKI KAISHA) 2 December 1992 * page 3, line 36 - line 43 * * page 3, line 49 - page 4, line 8 * * page 10, line 5 - line 14 * * page 10, line 28 - line 47 * * page 11, line 2 - line 4 * ----- | 1-12 | |
| | | | TECHNICAL FIELDS SEARCHED (Int.Cl.6) |
| | | | C12P C12N |
| The present search report has been drawn up for all claims | | | |
| Place of search THE HAGUE | | Date of completion of the search 4 July 1995 | Examiner Montero Lopez, B |
| CATEGORY OF CITED DOCUMENTS | | | |
| X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document | | T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons * : member of the same patent family, corresponding document | |

EPO FORM 1503 (3.12.91) (POM/CM)

FIG. 1

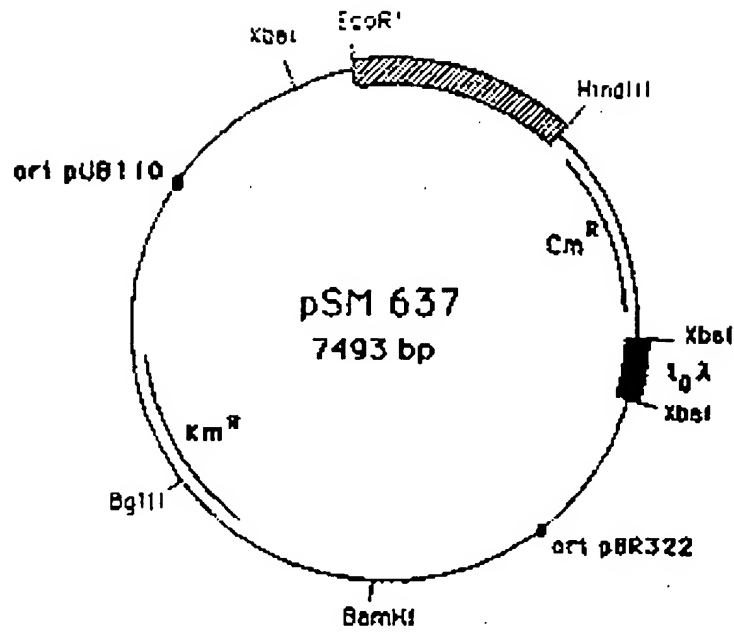


FIG. 2

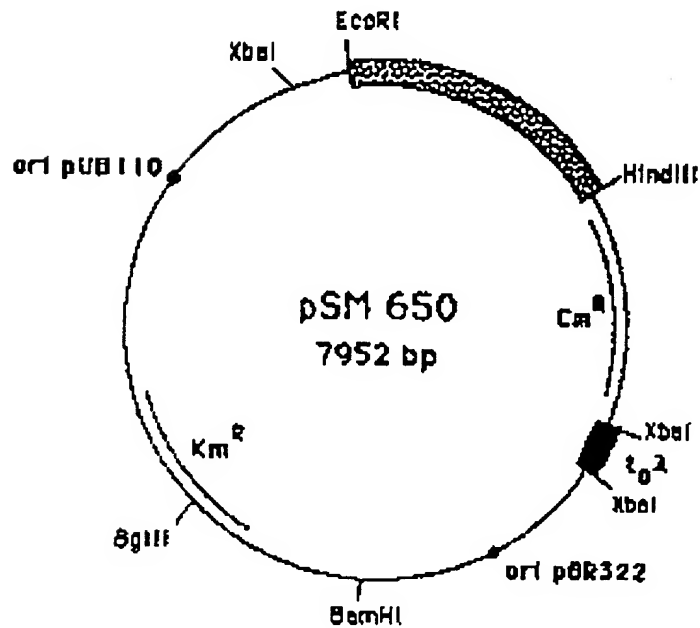


FIG. 3

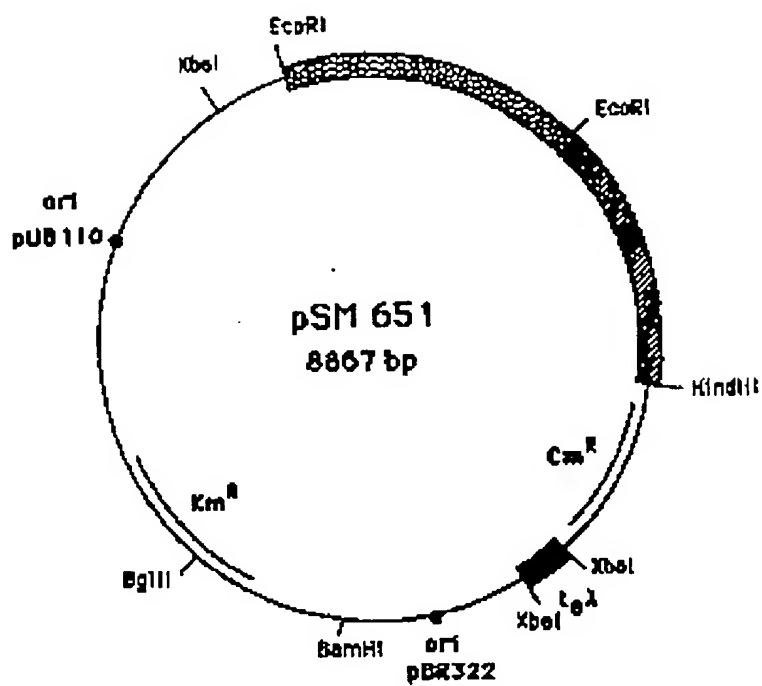


FIG. 4A

| | |
|---|-----|
| ATG ACA CGT CAG ATG ATA CTT GCT GTC CGA CAG CAA GGC CCC ATC | 45 |
| Met Thr Arg Gln Met Ile Leu Ala Val Gly Gln Gln Gly Pro Ile | |
| 5 10 15 | |
| CCG CGA GCG GAG ACA CCG GAA CAG GTC GTT GGC CCG CTC CTC GAC | 90 |
| Ala Arg Ala Glu Thr Arg Glu Gln Val Val Gly Arg Leu Leu Asp | |
| 20 25 30 | |
| ATG TTG ACG AAC GCA CCG AGC CCG GGC GTC AAC TTC ATC GTC TTT | 135 |
| Met Leu Thr Asn Ala Ala Ser Arg Gly Val Asn Phe Ile Val Phe | |
| 35 40 45 | |
| CCG GAG CTT GCG CTC ACG ACC TTC TTC CCG CCG TCG CAT TTC ACC | 180 |
| Pro Glu Leu Ala Leu Thr Thr Phe Phe Pro Arg Trp His Phe Thr | |
| 50 55 60 | |
| GAC GAG GCC GAG CTC CAT AGC TTC TAT GAG ACC GAA ATG CCC GGC | 225 |
| Asp Glu Ala Glu Leu Asp Ser Phe Tyr Glu Thr Glu Met Pro Gly | |
| 65 70 75 | |
| CCG GTC GTC CGT CCA CTC TTT GAG ACG GCG GCC GAA CTC GCG ATC | 270 |
| Pro Val Val Arg Pro Leu Phe Glu Thr Ala Ala Glu Leu Gly Ile | |
| 80 85 90 | |
| GCC TTC AAT CTG CCG TAC GCG GAA CTC GTC GTC GAA CCG GCG CTC | 315 |
| Gly Phe Asn Leu Gly Tyr Ala Glu Leu Val Val Glu Gly Gly Val | |
| 95 100 105 | |
| AAG CGT CCG TTC AAC ACG TCC ATT CTG GTC GAT AAG TCA CCG AAG | 360 |
| Lys Arg Arg Phe Asn Thr Ser Ile Leu Val Asp Lys Ser Gly Lys | |
| 110 115 120 | |
| ATC GTC CCG AAG TAT CGT AAG ATC CAT TTG CCG GGT CAC AAG GAG | 415 |
| Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro Gly His Lys Glu | |
| 125 130 135 | |
| TAC GAG GCC TAC CCG CCG TTC CAG CAT CTT GAA AAG CGT TAT TTC | 450 |
| Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys Arg Tyr Phe | |
| 140 145 150 | |

FIG. 4B

| | |
|---|-----|
| GAG CCG GGC GAT CTC GGC TTC CCG CTC TAT GAC GTC GAC GCC CCG | 495 |
| Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp Ala Ala | |
| 155 160 165 | |
| AAA ATG GGG ATG TTC ATC TGC AAC GAT CGC CGC TGG OCT GAA ACC | 540 |
| Lys Met Gly Met Phe Ile Cys Asn Asp Arg Arg Trp Pro Glu Thr | |
| 170 175 180 | |
| TGG CCG CTC ATG GGA CTT AAG GGC GCC GAG ATC ATC TGC GGC GGC | 585 |
| Trp Arg Val Met Gly Leu Lys Gly Ala Glu Ile Ile Cys Gly Gly | |
| 185 190 195 | |
| TAC AAC ACC CCG ACC CAC AAT CCG CCG GTT CCG CAG CAC CAC CAT | 630 |
| Tyr Asn Thr Pro Thr His Asn Pro Pro Val Pro Gln His Asp His | |
| 200 205 210 | |
| CTG ACC TCC TTC CAC CAC CTT CTG TCG ATG CAG GCC GGG TCG TAC | 675 |
| Leu Thr Ser Phe His His Leu Leu Ser Met Gln Ala Gly Ser Tyr | |
| 215 220 225 | |
| CAA AAC GGC GCC TCG TCC GCG GCG GCG GCG AAG CTC GCC ATG CAG | 720 |
| Gln Asn Gly Ala Trp Ser Ala Ala Ala Gly Lys Val Gly Met Glu | |
| 230 235 240 | |
| GAG CCG TCC ATG CTG CTC GGC CAT TCG TCC ATC CTG CCG CCG ACC | 765 |
| Glu Gly Cys Met Leu Leu Gly His Ser Cys Ile Val Ala Pro Thr | |
| 245 250 255 | |
| GGC GAA ATC GTT GCC CTG ACC ACC ACC TTG GAA GAC CAG CTG ATC | 810 |
| Gly Glu Ile Val Ala Leu Thr Thr Thr Leu Glu Asp Glu Val Ile | |
| 260 265 270 | |
| ACC GCC GCC CTC GAT CTC GAC CCG TGC CCG GAA CTG CCG GAA CAC | 855 |
| Thr Ala Ala Val Asp Leu Asp Arg Cys Arg Glu Leu Arg Glu His | |
| 275 280 285 | |
| ATC TTC AAT TTC AAA GCC CAT CGT CAG CCA CAG CAC TAC GGT CTG | 900 |
| Ile Phe Asn Phe Lys Ala His Arg Gln Pro Gln His Tyr Gly Leu | |
| 290 295 300 | |
| ATC GCG GAA TTT TGA | 915 |
| Ile Al Glu Phe STOP | |

FIG. 5A

| | |
|---|-----|
| ATG GAT ATC ATC ATC AAG AAC GGA ACC ATC GTA ACC GCG GAC GCG | 45 |
| Met Asp Ile Ile Ile Lys Asn Gly Thr Ile Val Thr Ala Asp Gly | |
| 5 10 15 | |
| ATT TCT CCC GCG GAT CTC GGA ATC AAG GAT GGC AAG ATC GCG CAG | 90 |
| Ile Ser Pro Ala Asp Leu Gly Ile Lys Asp Gly Lys Ile Ala Gln | |
| 20 25 30 | |
| ATC GCG GGA ACG TTC GCG CCG GCG GCG CCG ACA ATC CAC GCG TCC | 135 |
| Ile Gly Gly Thr Phe Gly Pro Ala Gly Arg Thr Ile Asp Ala Ser | |
| 35 40 45 | |
| GCG GCG TAC GTT TTT CCG GCG GCG ATC GAC GTT CAT ACG CAT GTC | 180 |
| Gly Arg Tyr Val Phe Pro Gly Gly Ile Asp Val His Thr His Val | |
| 50 55 60 | |
| GAG ACG GTC ACG TTC AAC ACG CAG TCG GCG GAC ACA TTC GCA ACC | 225 |
| Glu Thr Val Ser Phe Asn Thr Gln Ser Ala Asp Thr Phe Ala Thr | |
| 65 70 75 | |
| GCG ACG GTC GCG GCG GCG TGT GCG GCG ACG ACG ACC ATC GTC GAT | 270 |
| Ala Thr Val Ala Ala Ala Cys Gly Gly Thr Thr Thr Ile Val Asp | |
| 80 85 90 | |
| TTC TGC CAG CAG GAC CCG GCG CAT ACG CTC ACG GAG CCG GTC GCG | 315 |
| Phe Cys Gln Gln Asp Arg Gly His Ser Leu Arg Glu Ala Val Ala | |
| 95 100 105 | |
| AAA TCG GAC GCG ATG GCG GCG GCG AAG TCG GCG ATC GAC TAC GCG | 360 |
| Lys Trp Asp Gly Met Ala Gly Gly Lys Ser Ala Ile Asp Tyr Gly | |
| 110 115 120 | |
| TAC CAT ATC ATC GTG CTC GAT CCG ACT GAT ACG GTG ATC GAG GAG | 405 |
| Tyr His Ile Ile Val Leu Asp Pro Thr Asp Ser Val Ile Glu Glu | |
| 125 130 135 | |
| CTA GAG GTA CTG CCA GAT CTC GCG ATC ACG TCC TTC AAG GTC TTC | 450 |
| Leu Glu Val Leu Pro Asp Leu Gly Ile Thr Ser Phe Lys Val Phe | |

FIG. 58

| | | | | |
|---|-----|-----|-----|-----|
| | 140 | 145 | 150 | |
| ATC GCT TAT CCG GGC ATG AAC ATG ATC CAC GAC CTC ACA CTG CTC | | | | 495 |
| Met Ala Tyr Arg Gly Met Asn Met Ile Asp Asp Val Thr Leu Leu | | | | |
| | 155 | 160 | 165 | |
| AGG ACC CTC CAC AAG GCG GCG AAG ACT GGG TCA CTC GTC ATG GTC | | | | 540 |
| Arg Thr Leu Asp Lys Ala Ala Lys Thr Gly Ser Leu Val Met Val | | | | |
| | 170 | 175 | 180 | |
| CAC GCG GAG AAC GCG CAC GCG GCG GAC TAT CTT CCG GAC AAG TTC | | | | 585 |
| His Ala Glu Asn Gly Asp Ala Ala Asp Tyr Leu Arg Asp Lys Phe | | | | |
| | 185 | 190 | 195 | |
| CTC GCG GAT GCG AAA ACC GCG CCG ATC TAC CAC GCG CTC ACC CGT | | | | 630 |
| Val Ala Asp Gly Lys Thr Ala Pro Ile Tyr His Ala Leu Ser Arg | | | | |
| | 200 | 205 | 210 | |
| CCG CCG CCG GTC GAA CCG GAG GCG ACC GCG AGG GCG CTC GCG CTG | | | | 675 |
| Pro Pro Arg Val Glu Ala Glu Ala Thr Ala Arg Ala Leu Ala Leu | | | | |
| | 215 | 220 | 225 | |
| GCG GAA ATC CTC AAC CCG CCG ATC TAC ATC GTG CAT CTC ACC TGC | | | | 720 |
| Ala Glu Ile Val Asn Ala Pro Ile Tyr Ile Val His Leu Thr Cys | | | | |
| | 230 | 235 | 240 | |
| GAA GAA TCC TTC GAC GAG TTG ATG CCG CCA AAG GCT CCG GGT GTC | | | | 765 |
| Glu Glu Ser Phe Asp Glu Leu Met Arg Ala Lys Ala Arg Gly Val | | | | |
| | 245 | 250 | 255 | |
| CAC GCG CTC CCG GAA ACC TCG ACA CAA TAC CTC TAC CTC ACC AAG | | | | 810 |
| His Ala Leu Ala Glu Thr Cys Thr Gln Tyr Leu Tyr Leu Thr Lys | | | | |
| | 260 | 265 | 270 | |
| GAC GAC CTC GAG CCG CCG GAT TTC GAG GCG GCG AAG TAT GTT TTC | | | | 855 |
| Asp Asp Leu Glu Arg Pro Asp Phe Glu Gly Ala Lys Tyr Val Phe | | | | |
| | 275 | 280 | 285 | |
| ACC CCG CCT CCG CCG ACC AAG AAG GAC CAG GAA ATC CTC TGG AAC | | | | 900 |
| Thr Pro Pro Pro Arg Thr Lys Lys Asp Gln Glu Ile Leu Trp Asn | | | | |
| | 290 | 295 | 300 | |
| GCA CTC CCG AAC GCG CTC CTC GAA ACC GTC TCG TCG GAC CAT TGT | | | | 945 |
| Ala Leu Arg Asn Gly Val Leu Glu Thr Val Ser Ser Asp His Cys | | | | |

FIG. 5C

| | | | | |
|-----------------|---|-----|-----|------|
| | 305 | 310 | 315 | |
| TCC TGG CTC TTC | GAG GGG CAC AAG GAT CGG GGC AGG AAC GAC TTC | | | 990 |
| Ser Trp Leu Phe | Glu Gly His Lys Asp Arg Gly Arg Asn Asp Phe | | | |
| | 320 | 325 | 330 | |
| CCG GCC ATC CCG | AAC GGA GGG CCG GGC GTC GAG GAG CCG CTG ATG | | | 1035 |
| Arg Ala Ile Pro | Asn Gly Ala Pro Gly Val Glu Glu Arg Leu Met | | | |
| | 335 | 340 | 345 | |
| ATG GTC TAT CAG | GGC GTC AAC GAA GGC CGC ATT TCC CTC ACC CAG | | | 1080 |
| Met Val Tyr Gln | Gly Val Asn Glu Gly Arg Ile Ser Leu Thr Gln | | | |
| | 350 | 355 | 360 | |
| TTC GTA GAA CTC | GTC GGC ACG CGC CCG GGC AAC GTC TTC GGC ATG | | | 1125 |
| Phe Val Glu Leu | Val Ala Thr Arg Pro Ala Lys Val Phe Gly Met | | | |
| | 365 | 370 | 375 | |
| TTC CCG GAA AAA | GGA ACG GTC GCG GTC GGT TCG GAT GCC GAC ATC | | | 1170 |
| Phe Pro Glu Lys | Gly Thr Val Ala Val Gly Ser Asp Ala Asp Ile | | | |
| | 380 | 385 | 390 | |
| GTC CTT TCG GAT | CCG GAG GCT GAA ATG CTC ATC GAA CAA AGC GCG | | | 1215 |
| Val Leu Trp Asp | Pro Glu Ala Glu Met Val Ile Glu Gln Ser Ala | | | |
| | 395 | 400 | 405 | |
| ATG CAT AAC GCC | ATG GAT TAC TCC TCC TAC GAG GGA CAC AAG ATC | | | 1260 |
| Met His Asn Ala | Met Asp Tyr Ser Ser Tyr Glu Gly His Lys Ile | | | |
| | 410 | 415 | 420 | |
| AAG GGC GTG CCG | AAG ACA GTG CTG CTG CGT GGC AAG GTG ATC GTC | | | 1305 |
| Lys Gly Val Pro | Lys Thr Val Leu Leu Arg Gly Lys Val Ile Val | | | |
| | 425 | 430 | 435 | |
| GAC GAG GGA ACC | TAC GTG GGG GCG CCG ACG GAT GGC CAG TTC CCG | | | 1350 |
| Asp Glu Gly Thr | Tyr Val Gly Ala Pro Thr Asp Gly Gln Phe Arg | | | |
| | 440 | 445 | 450 | |
| AAG GGC GGC AAA | TAC AAG CAA TAA | | | 1373 |
| Lys Arg Arg Lys | Tyr Lys Gln STOP | | | |
| | 455 | | | |

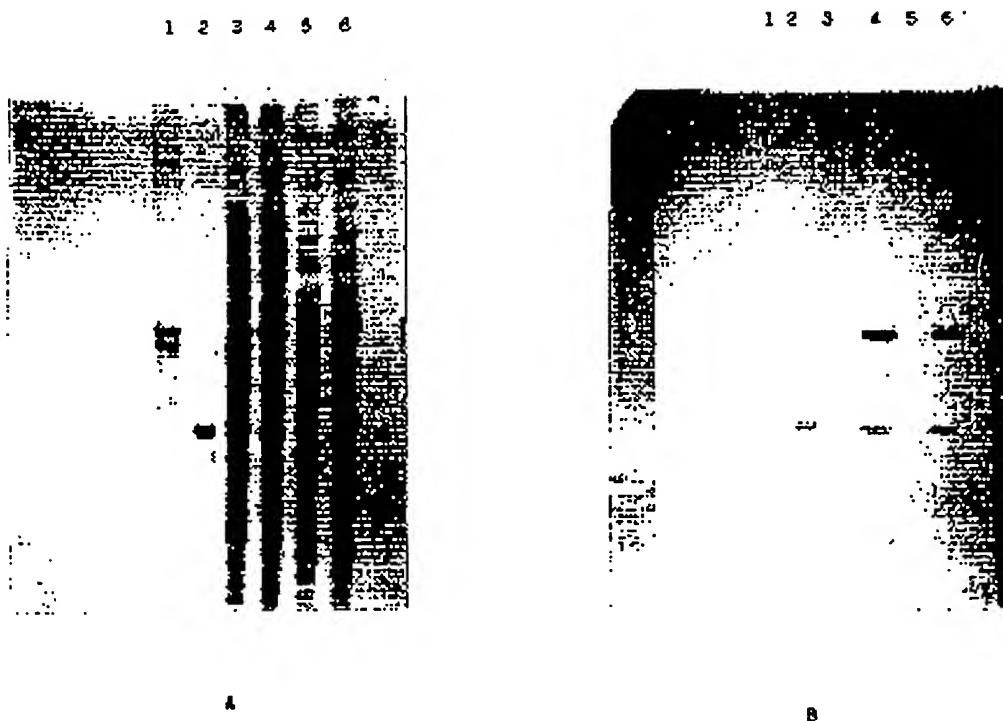


FIG. 6